SC EDUCATION OVERSIGHT COMMITTEE

Reporting facts. Measuring change. Promoting progress.

AGENDA

Academic Standards and Assessments Subcommittee Meeting
Monday, September 17, 2007
10:00 AM
Room 215, Blatt Building

I.	Welcome and Introductions	Mr. Tom DeLoach
II.	Approval of the May 21, 2007 Meeting Minutes	Mr. Tom DeLoach
III.	Approval: Review of US History End of Course Test	Mr. David Potter
IV.	Approval: Review of SC-Alternative ELA and Mathematics Assessments	Mr. David Potter
V.	Continued Discussion: Cyclical Review of PACT ELA and Math Assessments	Ms. Elizabeth Jones Mr. David Potter
VI.	Other Business	Mr. Tom DeLoach
Adjournment		Mr. Tom DeLoach

<u>Academic Standards and Assessments Subcommittee Members:</u>

Mr. Tom DeLoach, Chairman
Sen. Mike Fair
Sen. Wes Hayes
Mrs. Buffy Murphy
Rep. Joe Neal
Rep. Bob Walker
Dr. Kristi Woodall

SOUTH CAROLINA EDUCATION OVERSIGHT COMMITTEE
Subcommittee on Academic Standards and Assessments

Minutes of the Meeting
May 21, 2007
10:00 AM, Room 215 Blatt Building

Subcommittee Members Present: Tom DeLoach, Wes Hayes, Buffy Murphy, Robert Walker, Kristi Woodall
Staff Present: Jo Anne Anderson, Paul Horne, David Potter

Welcome and Introductions
Mr. DeLoach welcomed members and guests to the meeting.

Minutes
The minutes from the March 19, 2007 meeting were approved as written.

Information: Cyclical Review of the ELA Standards
Dr. Paul Horne reviewed the *Procedures for the Cyclical Review of Current South Carolina K-12 Academic Standards and for the Development of New Academic Standards* and the application of those procedures to the cyclical review of the ELA and Mathematics Academic Standards. Subcommittee members discussed the application of the procedures and how the procedures might be adapted to provide flexibility to the standards review process.

Approval: Cyclical Review of the Mathematics Standards
Dr. Paul Horne presented the revised South Carolina Mathematics Academic Standards for approval. The Subcommittee had previously approved the standards at its March 2007 meeting and the full EOC subsequently approved the standards at its April 2007 meeting. However, at its April 2007 meeting the State Board of Education amended the Mathematics standards. One amendment was made to an indicator at the grade 3 level and the other was made to a grade 4 indicator. These amended standards were then considered by the Subcommittee at its May 2007 meeting. Following discussion of the impact of the amendments, the Subcommittee adopted the recommendation that the standards be approved as amended and submitted to the full EOC for approval.

Approval: Use of End of Course Test Scores in School Ratings
David Potter presented a set of recommendations which provide for the use of End of Course test data in the middle school Absolute Ratings, clarify the attribution of End of Course test scores from the Virtual High School and dual credit courses for reporting and accountability purposes, and provide for the reporting and use of End of Course test results and school profile data from schools containing grade 9 only. The high school End of Course test results are currently included in the calculation of high school and school district ratings, but are not included in the calculation of middle school ratings. The proposed recommendations were reviewed by educators in the field prior to consideration by the Subcommittee; the educators' responses were reviewed by the Subcommittee in its consideration of the proposed recommendations. The recommendations would take effect with the 2007-2008 school year. The Subcommittee discussed the recommendations and adopted them for consideration by the full EOC.

Approval: Cyclical Review of PACT ELA and Mathematics Assessments

As it does for curriculum standards, the EAA in Section 59-18-360 (A) also establishes a cyclical review at least every seven years of the standards based assessments. David Potter presented the results and recommendations from the cyclical review of the PACT ELA and Mathematics assessments conducted by EOC and SDE staff. The Subcommittee discussed the results and proposed recommendations from the cyclical review, but took no action.

There being no further business, the subcommittee adjourned.

EDUCATION OVERSIGHT COMMITTEE

Subcommittee: Academic Standards and Assessments

Date: September 17, 2007

REPORT/RECOMMENDATION

Review of the U.S. History and the Constitution End of Course Field Test

PURPOSE/AUTHORITY

Section 59-18-320. (A) After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations.
Section 59-18-320 (C). After review and approval by the Education Oversight Committee, the end of course assessments of benchmark courses will be administered to all public school students as they complete each benchmark course.

CRITICAL FACTS

The U.S. History and the Constitution end of course field test was administered Spring 2006. Recommendations regarding the test must be communicated to the SC State Department of Education, which must respond within one month. State assessments must be reviewed and approved by the Education Oversight Committee.

TIMELINE/REVIEW PROCESS

The U.S. History and the Constitution field test was reviewed by the EOC in December 2006 and recommended that the test continue as a field test and that teachers be surveyed regarding their coverage of the course standards in instruction. The survey results indicated that teachers did not have sufficient time to cover all the standards adequately and were not teaching all of the standards.

ECONOMIC IMPACT

Cost:

Fund/Source:

ACTION REQUEST

☒ **For approval** ☐ **For information**

ACTION TAKEN

☐ **Approved** ☐ **Amended**

☐ **Not Approved** ☐ **Action deferred (explain)**

SC EDUCATION OVERSIGHT COMMITTEE

Reporting facts. Measuring change. Promoting progress.

MEMORANDUM

TO: Members of the Academic Standards and Assessments Subcommittee
Tom DeLoach, Chairman
Mike Fair
Wes Hayes
Buffy Murphy
Joe Neal
Bob Walker
Kristi Woodall

FROM: David Potter
Director of Research

DATE: September 4, 2007

Subject: Review of the U.S. History and the Constitution End of Course Field Test

The Education Oversight Committee (EOC) is charged in the Education Accountability Act to review the field tests for new assessments in the state assessment program:

> "After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations." (Section 59-18-320 A)

The EOC also has a role in the adoption of state assessments:

> "Any new standards and assessments required to be developed and adopted by the State Board of Education, through the Department of Education, must be developed and adopted upon the advice and consent of the Education Oversight Committee." (Section 59-18-320 D)

In December 2006 the EOC reviewed the U.S. History and the Constitution End of Course field test and adopted the following recommendations:

1. Continue the field test of the U.S. History and the Constitution end of course test during 2006-2007 by administering the currently prepared draft operational forms to students enrolled in the course. Monitor the performance of students on the U.S. History and the Constitution tests administered in the 2006-2007 school year and evaluate the technical characteristics of the items and the performance standards in Summer 2007 for possible revision.
2. In cooperation with the State Department of Education, survey U.S. History and the Constitution teachers in Spring 2007 to describe their understanding and use of the U.S. History and the Constitution standards and relate the results to student performance.

In April and May 2007 a survey regarding the course academic standards was distributed to the 633 teachers of the U.S. History and the Constitution course identified by 84 of the 85 school districts in South Carolina. A total of 312 teachers (49.3%) responded to all the survey questions. The teachers' responses to the survey were summarized and distributed to EOC members at their August 13, 2007 meeting. In addition, the survey results were distributed to the members of the Instructional Leaders' Roundtable and EOC staff discussed the findings with them at their August 16, 2007 meeting.

<u>Strengths identified in the U.S. History and the Constitution End of Course test:</u>

- The test is well-aligned to the academic course standards;
- The cognitive demands of the test items are well-aligned with the rigorous academic course standards;
- The test has adequate technical characteristics, although many of the test items are difficult (the average score on the field test was 41% correct);
- The test can provide a measure of accountability for implementation of high academic standards.

<u>Concerns Identified:</u>

The difficulty of the test diminishes its effectiveness to "differentiate levels of achievement" (Section 59-18-320 A). The survey of U.S. History and the Constitution teachers in May 2007 and subsequent discussions with members of the Instructional Leaders' Roundtable suggest the following factors accounting for the low student achievement observed:

- There is poor alignment between the content of the standards and the content of classroom instruction;
- Teachers reported several factors related to time which adversely affect students' learning of the course standards:
  - ✓ There is too little time to teach all of the standards, especially in one-semester "block" classes;
  - ✓ There is also too little time to teach for students to attain the higher levels of understanding required for the course standards and the test;
  - ✓ Some teachers report concerns that the test is administered too soon before the end of the semester to complete instruction on all of the standards;

- Teachers reported they need help with determining an effective pace for teaching the course standards, especially when time is limited;
- Teachers reported that support materials for professional development are not available or are inadequate;
- Students may not be motivated to perform well on the tests when there are no perceived consequences to them for poor performance or benefits for high performance.

Recommendations:

1. The State Department of Education (SDE) should take actions to improve the alignment among the U.S. History and the Constitution course standards, the instruction of those standards, and the End of Course test. Prior to EOC approval the SDE should provide evidence for the enactment of those actions to the EOC. The actions to improve the alignment may include, in addition to other possible activities:
   - ✓ Examine the course standards and End of Course test to identify or affirm the essential content to be learned and tested;
   - ✓ Complete the development of the Teacher's Guide, including guides for effectively pacing instruction, to the U.S. History and the Constitution course standards and End of Course test.

2. Continue the administration of the U.S. History and the Constitution End of Course test as a field test and provide feedback to schools and districts on the performance of their students.

3. The actions undertaken to improve the alignment among the standards, instruction, and the test should be accomplished by June 2008 to allow for professional development activities with teachers during Summer 2008.

As specified in Section 59-18-320 A, the SDE must respond to recommendations regarding the field test made by the EOC within one month. The EOC may approve the use of the test based upon the response from the SDE.

DRAFT

MEMORANDUM

TO: Members of the Academic Standards and Assessments Subcommittee, Education Oversight Committee

FROM: Teri Siskind
Deputy Superintendent for Curriculum Services and Assessment
SC Department of Education

David Potter
Director of Research
Division of Accountability, Education Oversight Committee

DATE: May 21, 2007

SUBJECT: Cyclical Review of PACT ELA and Mathematics Assessments

The Education Accountability Act of 1998 (EAA) provides for the establishment of tests based on the state academic standards in English Language Arts (ELA), Mathematics, Science and Social Studies (Section 59-18-310). The EAA in Section 59-18-360 (A) also establishes a cyclical review every seven years of the academic standards and the standards based assessments: "The State Board of Education, in consultation with the Education Oversight Committee, shall provide for a cyclical review by academic area of the state standards and assessments to ensure that the standards and assessments are maintaining high expectations for learning and teaching. All academic areas must be initially reviewed by the year 2005. At a minimum, each academic area should be reviewed and updated every seven years. After each academic area is reviewed, a report on the recommended revisions must be presented to the Education Oversight Committee for its consideration. After approval by the Education Oversight Committee, the recommendations may be implemented. As a part of the review, a task force of parents, business and industry persons, community leaders, and educators, to include special education teachers, shall examine the standards and assessment system to determine rigor and relevancy." This memorandum and the attached materials present the results and recommendations from the cyclical review of the Palmetto Challenge Achievement Test (PACT) assessments in ELA and Mathematics.

The content for the cyclical review of PACT ELA and Mathematics was established through several meetings of EOC and State Department of Education (SDE) staff members. Many of the issues studied for the review were identified in the February 2005 report of the Testing Task Force. The following issues were studied for the review:

1. Study the elimination of PACT Mathematics and ELA, grades 1 and 2;
2. Study the feasibility of equating Algebra 1 and English 1 End of Course tests to grade level PACT in Mathematics and ELA;
3. Review the Advanced and Proficient levels of PACT Mathematics and ELA;
4. Study design of score reports to improve reporting of PACT ELA and Mathematics results;
5. Study the feasibility of reporting PACT ELA and Mathematics results at the strand level;
6. Study the effects of the oral administration of grade 3 tests on performance;
7. Study the feasibility of vertically equating PACT Mathematics and ELA scores;

8. Conduct a controlled cost and program effectiveness study of the State's readiness for online administration of PACT Mathematics and ELA;
9. Study the elimination of Constructed Response items on PACT Mathematics and ELA.

All but one of the studies listed above were conducted either by SDE staff or by outside consultants or organizations, including testing contractors. The review of PACT ELA and Mathematics Proficient and Advanced standards (number 3 in the list of issues above) was conducted by EOC staff. A report summarizing the findings of the studies is attached in the Appendix to this memorandum. Copies of the studies are available on the Education Oversight Committee web site (eoc.sc.gov).

Recommendation 1

Revise the current PACT ELA "Advanced" cut scores to better distinguish performance at the two higher performance levels. This recommendation should be implemented for the Spring 2008 administration of the PACT ELA tests. Revising only the Advanced cut score will not affect the percentages of students scoring at the Proficient level or higher.

This recommendation is based on a review of the PACT ELA assessment results and technical characteristics and on the recommendations of two groups which have studied the issue. New ELA assessments may be administered in Spring 2009 and new standards would be set on the tests following that administration. The current PACT would continue in use for the 2007 and 2008 administrations. Thus, revisions to the Advanced cut scores on the PACT ELA tests may apply only to scores from a single administration - the 2008 administration. The implementation of this recommendation may not be cost effective if the PACT ELA tests are administered for only one more year.

- The empirical review of statewide data found that over the seven years 1999-2005 the average percentage of students scoring Advanced in ELA ranged from 1.9% in grade 5 to 5.3% in grade 3. This contrasts with PACT Mathematics results over the same period, in which the percent of students Advanced ranged from 6.6% in grade 8 to 11.4% in grade 4. The average number of students statewide scoring Advanced in ELA across the years ranged from 918 in grade 5 to 2,596 in grade 3, compared to a range in Mathematics from 3,202 in grade 8 to 5,695 in grade 4. While one might not expect the ELA and Mathematics percentages of students Advanced to be identical, there is a wide gap between the two subject areas and it difficult for students to score at the Advanced level in ELA.
- A review of the extent to which students who score at the Advanced level in ELA and Mathematics maintain that performance level in the subsequent year of testing was conducted based on longitudinally matched data. The review found that students initially scoring Advanced in Mathematics were more likely to score Advanced on the following year's test than students scoring Advanced in ELA. There was also wider variation between grade levels in ELA than in Mathematics in the percentages of students who maintain their Advanced performance level from year to year. This suggests that the Advanced cut scores are not as well aligned from grade to grade in ELA as in Mathematics.
- A review of the Conditional Standard Errors of Measurement (CSEM) at the Advanced performance levels for ELA and Mathematics found that the largest CSEM over the six years between 2000 and 2005 was at grade 5 in ELA, and that the CSEMs in ELA were

generally larger in ELA than in Mathematics (Tables 3 and 4). The CSEM provides an index of the reliability or accuracy of a given score on a test: the smaller the CSEM for a test score, the more reliable that score is. This review suggests that the Advanced cut scores in ELA are less reliable than those in Mathematics.

- After reviewing data at their February 2007 meeting, the National Technical Advisory Committee to the EOC Division of Accountability indicated that the Advanced cut point in PACT ELA should be studied, noting its extreme difficulty.
- The Testing Task Force, in its February 2005 final report, also expressed concerns about the ELA Advanced scores, stating in its recommendations that, "Currently, the Advanced level of the test needs the most attention, particularly in English Language Arts" (p. 9, Testing Task Force, 2005).

Recommendation 2

Develop or adopt new standards based assessments of ELA and Mathematics for grades 3 through 8 to replace the current PACT. The new tests must meet the criteria for technical quality, proficiency level expectations, and reporting called for in the Education Accountability Act and No Child Left Behind and must be designed to provide appropriate information to meet the requirements of the state and federal accountability systems. The design of the new tests should facilitate the measurement of student growth over time as well as student performance at the end of each grade level. The new tests should reflect improvements and current best practices in test design, administration, and reporting of results, including the appropriate use of technology. The new tests should first be administered no later than Spring 2011.

This recommendation is based on the following findings:

- This is an appropriate time to make changes to the testing program.
  - ✓ PACT ELA and Mathematics tests were first administered in 1999. Spring 2007 testing marks their ninth year of use, and, if PACT is not replaced by then, by 2011 the tests will have been administered for 13 years. Because of changes to academic content standards and concerns about security of items which have been in use over an extended period of time, state tests usually are changed by the time they reach the age of the current PACT ELA and Mathematics tests.
  - ✓ The adoption of revised academic standards for ELA and Mathematics in 2007 will require the revision of the standards based tests aligned with those standards, so this is a good time to develop or adopt new tests based on those revised academic standards.
  - ✓ The current adoption by states of more extensive use of technology such as computers and the internet in their testing programs for administration, scoring, and reporting represents the future of large scale assessment. As stated by the Testing Task Force in its 2005 report, "The future of assessment is computerized. The state should position itself to administer and score all assessments electronically" (page 3). A study of computer-based testing is currently underway and recommendations will be made in June.

- Development or adoption of new tests is needed to deal with the perceived shortcomings of the current PACT ELA and Mathematics tests.
  - ✓ Reporting of results: The cyclical review of reports on studies to provide strand-level information from PACT ELA and Mathematics indicate that the tests are not designed to provide this information and thus there is insufficient information for some strands to provide accurate measures of strand-level performance. The EAA requires that the accountability assessment results be reported in a manner that is useful for curriculum review and adjustment of instruction and in a format easily understood by families, educators, and the public. The EAA directs, "The Department of Education is directed to provide assessment results annually on individual students and schools in a manner and format that is easily understood by parents and the public. In addition, the school assessment results must be presented in a format easily understood by the faculty and in a manner that is useful for curriculum review and instructional improvement." (Section 59-18-370). The new tests should be designed to provide useful information for instructional use by teachers and administrators and for evaluation by parents of their children's achievement.
  - ✓ Measurement of student growth in achievement: The state report card Improvement Rating is required by the EAA to be based on individual student growth relative to the academic standards based on longitudinally matched test data. Currently, the federal government is investigating the use of student academic growth models for calculating Adequate Yearly Progress (AYP). The reports on studies of vertically scaling PACT ELA and Mathematics reviewed for the cyclical review indicate that the PACT tests cannot provide a vertical scale of sufficient reliability at the scale score level to make accurate evaluations of individual student growth in achievement from year to year. However, ELA and Mathematics tests which support accurate and meaningful evaluations of student growth would be desirable both for instructional and accountability use. In developing new tests a solution should be sought to the problem of establishing meaningful growth measures.

The PACT ELA and Mathematics tests have served their purpose well as an accountability measure but have limitations with respect to the reporting of strand level results and the measurement of student academic growth. These concerns coupled with the adoption of new academic content standards make this an opportune time to move testing in South Carolina to the next generation of testing.

APPENDIX
Cyclical Review of PACT English Language Arts And Mathematics Assessments
May 2007

Section 59-18-320 of the Education Accountability Act charges the Education Oversight Committee (EOC) with the review of the assessments in the statewide assessment program for alignment, level of difficulty and validity, and for the ability to differentiate levels of achievement. Section 59-18-360 indicates that the State Board of Education, in consultation with the Education Oversight Committee, shall provide for a cyclical review of assessments. The State Department of Education and the Accountability Division of the EOC conducted a series of meetings and determined that the cyclical review of assessments would address the areas presented in this report. Reports cited in this review are listed at the end of the section and are available on the Education Oversight Committee web site (eoc.sc.gov).

Elimination of PACT mathematics and ELA for grades 1 and 2

The Palmetto Achievement Challenge Tests (PACT) in mathematics and English language arts (ELA) for grades 1 and 2 were originally developed for use by districts. A state proviso established off-grade-level testing which permitted students with Individual Education Programs (IEPs) to take tests consonant with their instructional level. Hence, an eighth grade student who was being instructed at a fifth grade level was tested with a fifth grade test if designated by his IEP team.

Since South Carolina had tests in mathematics and ELA for grades 1 and 2, the state permitted the use of these tests as off-grade-level tests even though the lowest grade level in the statewide assessment program was grade 3. Over time, the state development cycle included these off-grade-level tests in grades 1 and 2, used only for a small number of students.

Several initiatives converged to result in the elimination of PACT in grades 1 and 2. There was growing concern about the extension of off-grade-level testing beyond the parameters of the statewide testing program. Many off-grade-level tested students took the grades 1 and 2 tests even when they reached middle-school age.

Not only was the context not age-appropriate, there was concern about the content expectations established by these lower-grade level tests.

In addition, the Testing Task Force established in 2004 was charged with reviewing the costs of the testing program. Off-grade-level testing for grades 1 and 2 cost about $500,000 per year. Simultaneously, through provisions of No Child Left Behind, off-grade-level testing was being discouraged. Students tested off-grade-level could not be considered as tested for accountability purposes. After a two-year phase out, off-grade-level testing was discontinued for the 2006–07 school year.

Equating Algebra 1 and English 1 end-of-course tests to PACT

Students taking Algebra 1 and English 1 for credit are required to participate in the end-of-course examination program (EOCEP). Some students in middle school (predominantly eighth grade) take these courses and the required examinations. EOCEP is administered at the end of the course and PACT is administered at the end of the school year; therefore, some middle school students were required to participate in both testing programs within a few weeks. In an attempt to reduce the amount of testing, the Testing Task Force recommended that the Department review the tests to determine whether EOCEP and PACT could be "equated" for these two subjects.

The attached reports, "Investigating the Projection of PACT Scores from EOCEP," are based on 2004 test data. The reports were presented to the South Carolina Technical Advisory Committee in 2005 along with 2005 test data. Although the TAC "expressed some sentiment to provide relief to middle school students who are double tested" (TAC Proceedings 2005), there was no statistical justification for projecting PACT scores on the basis of EOCEP alone. There was also some concern for "equating" the subject matter of the corollary tests.

Reports:
Investigating the Projection of PACT Scores from EOCEP (ELA), 2005
Investigating the Projection of PACT Scores from EOCEP (Math), 2005

Advanced and Proficient Levels on PACT

The Testing Task Force, in its February 2005 final report, expressed concerns about the Proficient and Advanced performance cut scores on PACT ELA and Math. The group was particularly concerned about the ELA Advanced scores, stating in its recommendations that, "Currently, the Advanced level of the test needs the most attention, particularly in English Language Arts" (p. 9, Testing Task Force, 2005).

The review of PACT ELA and Mathematics Proficient and Advanced standards was conducted by EOC staff. The results of that review are presented in their entirety in this document.

Several independent studies have found the PACT ELA and Mathematics Proficiency performance standards to be generally well aligned with the performance standards

of the National Assessment of Educational Progress (NAEP). South Carolina has been recognized for these high ELA and Mathematics proficiency standards in national studies reported in the journal *Education Next* and by the group The Education Trust.

The relationships between PACT ELA performance standards and NAEP Reading performance standards are illustrated in Figures 1 – 4. Figures 1 and 2 show the percentages of South Carolina 4th graders (Figure 1) and 8th graders (Figure 2) who scored at the Proficient level or higher on NAEP and PACT for the years 2002 through 2005 (the most recent year data are available). The performance of all students nationally on NAEP is also shown on all figures to provide a comparison. At both the 4th and 8th grade levels the performance at the Proficient level of South Carolina students on NAEP Reading was lower than all students nationally. At the 8th grade the percentages of South Carolina students scoring Proficient or higher on PACT ELA was similar to the performance of students nationally on NAEP, but at the 4th grade level South Carolina students consistently scored higher on PACT than they did on NAEP, suggesting that the PACT Proficient standard was somewhat lower than the NAEP standard. Figures 3 and 4 illustrate the percentages of students who fail to meet the minimal performance standard (Basic) on NAEP and PACT. At the 8th grade (Figure 4) the percentages of South Carolina students who fail to meet the Basic standard are similar on both PACT and NAEP, but at the 4th grade (Figure 3) many more South Carolina students fail to meet the Basic standard on NAEP than on PACT, indicating that the grade 4 PACT Basic standard is easier than the NAEP Basic standard. This raises the question whether the grade 4 PACT is under-identifying students whose performance is low enough (Below Basic) that they need academic assistance in reading. However, the interpretation of these findings is complicated by the fact that NAEP is a test of reading, while PACT ELA tests reading, writing, and reference skills. Perhaps the relatively higher performance at the Basic level of 4th grade students on PACT compared to NAEP reading reflects their higher levels of writing and reference skills compared to their reading skills.

NAEP and PACT Mathematics performance is shown in Figures 5 – 8. The data in these figures indicate that PACT Mathematics Proficient and Basic performance standards at both the 4th and 8th grade levels are well aligned with NAEP Mathematics standards. A close alignment with NAEP performance expectations in

both ELA and Mathematics is necessary for SC to align its educational reforms to national standards and to measure its progress toward meeting those standards.

An empirical review of statewide PACT ELA and Mathematics data by Education Oversight Committee staff found that over the seven years 1999-2005 the average percentage of students scoring Advanced in ELA ranged from 1.9% in grade 5 to 5.3% in grade 3 (see Table 1). This contrasts with PACT Mathematics results over the same period, in which the percent Advanced ranged from 6.6% in grade 8 to 11.4% in grade 4 (Table 2). The average yearly number of students statewide scoring Advanced in ELA ranged from 918 in grade 5 to 2,596 in grade 3, compared to a range in Mathematics from 3,202 in grade 8 to 5,695 in grade 4. While one might not expect the ELA and Mathematics percentages of students Advanced to be identical, there is a wide gap between the two subject areas and it is extraordinarily difficult for students to score at the Advanced level in ELA.

Table 1
Average Percentages and Numbers Proficient or Advanced Over Seven Years
PACT ELA, Years 1999 Through 2005

Grade	Average Number Tested	Average Percent Proficient	Average Number Proficient	Average Percent Advanced	Average Number Advanced
3	48,693	38.7	18,837	5.3	2,596
4	49,157	32.2	15,780	2.6	1,284
5	49,394	24.4	12,033	1.9	918
6	49,916	24.2	12,070	4.9	2,456
7	49,922	22.5	11,251	3.1	1,565
8	48,561	20.9	10,147	3.7	1,808

Source: *Technical Documentation for the 2005 Palmetto Achievement Challenge Tests of English Language Arts, Mathematics, Science, and Social Studies,* SC Department of Education

Table 2

Average Percentages and Numbers Proficient or Advanced Over Seven Years

PACT Math, Years 1999 Through 2005

Grade	Average Number Tested	Average Percent Proficient	Average Number Proficient	Average Percent Advanced	Average Number Advanced
3	49,852	18.5	9,161	10.4	5,132
4	49,934	19.0	9,492	11.4	5,695
5	49,826	16.3	8,107	10.1	5,026
6	50,470	19.1	9,632	11.0	5,531
7	50,341	15.1	7,620	11.3	5,692
8	48,790	13.1	6,403	6.6	3,202

Source: *Technical Documentation for the 2005 Palmetto Achievement Challenge Tests of English Language Arts, Mathematics, Science, and Social Studies,* SC Department of Education

A review of the extent to which students who score at the Advanced level in ELA and Mathematics maintain that performance level in the subsequent year of testing was conducted based on longitudinally matched data. The review found that students initially scoring Advanced in Mathematics in most grade levels were more likely to score Advanced on the following year's test than students scoring Advanced in ELA (see Figures 9 and 10). There was also wider variation between grade levels in ELA than in Mathematics in the percentages of students who maintain their Advanced performance level from year to year. This suggests that the Advanced cut scores are not as well aligned from grade to grade in ELA as in Math.

A review of the Conditional Standard Errors of Measurement (CSEM) at the Advanced performance levels for ELA and Mathematics found that the largest CSEM over the six years between 2000 and 2005 was at grade 5 in ELA, and that the CSEMs in ELA were generally larger in ELA than in Mathematics(Tables 3 and 4). The CSEM provides an index of the reliability or accuracy of a given score on a test: the smaller the CSEM for a test score, the more reliable that score is. This review suggests that the Advanced cut scores in ELA are less reliable than those in Math.

Table 3

PACT ELA Conditional Standard Errors of Measurement (CSEM)

At Advanced Cut Score

Grade	1999	2000	2001	2002	2003	2004	2005
3	4.96	9.1	9.1	8.4	8.0	8.0	7.3
4	4.98	9.6	7.3	8.1	9.0	8.6	9.2
5	5.29	9.6	11.0	10.6	10.0	10.5	11.2
6	3.92	5.4	5.4	5.7	6.0	6.5	6.4
7	3.88	6.5	5.7	6.0	6.0	6.2	6.2
8	3.18	5.2	5.2	5.5	5.4	5.9	5.8

Source: Technical Documentation Reports for the 1999, 2000, 2001, 2002, 2003, and 2005 Palmetto Achievement Challenge Tests, SC Department of Education

Table 4

PACT Mathematics Conditional Standard Errors of Measurement (CSEM)

At Advanced Cut Score

Grade	1999	2000	2001	2002	2003	2004	2005
3	4.53	7.0	7.7	8.0	7.2	7.9	7.6
4	4.27	6.7	7.7	7.3	6.5	7.3	6.9
5	4.84	7.0	7.9	6.8	6.8	6.7	7.0
6	3.83	5.6	5.8	5.6	6.0	5.6	5.7
7	3.62	5.2	5.5	5.2	5.0	5.1	5.2
8	3.05	5.0	5.0	4.9	5.3	5.3	5.1

Source: Technical Documentation Reports for the 1999, 2000, 2001, 2002, 2003, and 2005 Palmetto Achievement Challenge Tests, SC Department of Education

After reviewing data at their February 2007 meeting, the National Technical Advisory Committee to the EOC Division of Accountability indicated that the Advanced cut point in PACT ELA should be studied and, if the test were to continue in use for several more years, should be revised as soon as possible.

Reference:

Final Report of SC Task Force on Testing, February 14, 2005.

Instructional Level Information for PACT

Designed as accountability measures, the Palmetto Achievement Challenge Tests were constructed to provide information about each student's performance in ELA, mathematics, science and social studies while minimizing test burden. PACT is administered at the end of the school year, based on requests from district instructional personnel. Though not timed, the typical completion is well under two hours per test day (ranging from 50 to 110 minutes depending upon the grade and subject area).

During the initial years of PACT administration, the state provided "instructional level information." After a couple of years, the Department became concerned about the accuracy of this information and discontinued its production with exceptions justified by data (Reading and Writing). Instructional level information is sought by educators, and the Department has conducted a number of studies over the years to determine whether and how this information could be provided. The analyses are summarized in two reports: "Strand Level Information from PACT Tests" and "An Analysis of Item Mapping and Test Reporting Strategies."

It is worthwhile to note that the Department does provide detailed Descriptions of Achievement Levels (see the report, *Performance-Level Descriptors for the Palmetto Achievement Challenge Tests,* 2005).

Reports:
An Analysis of Item Mapping and Test Reporting Strategies, 2003
Strand Level Information from PACT Tests, 2007

Oral Administration of PACT for Grade 3

Unlike other states, South Carolina has required teachers to read aloud the third grade tests for mathematics, science and social studies. While some educators advocate "read aloud" as appropriate for third grade children, other educators express concern about this practice. The Department studied the impact of "read aloud" as part of the 2006 field test administration and the results are summarized in "PACT Grade 3 "Read Aloud" Study."

Due to a change in law, the 2007 administration of PACT will sample students tested in science and social studies at all but two grades. Because approximately half of the third grade students within each school will take science tests and approximately half will take social studies tests, "read aloud" became a particular interest due to logistical concerns. While students in upper grades will not require separation for test administration of science and social studies tests, third grade students taking the science test cannot be tested in the same classroom as third grade students taking the social studies tests. In September, the director of the Office of Assessment polled test directors in other states about their practices of administration for grade 3. Twenty-six responses were received and represented 27 states and the Department of Defense Schools. None of the respondents requires or permits "read aloud" as a general practice.

Based on the findings of the study and the survey, the Department has determined that it will not discontinue "read aloud" at grade 3 until the testing program is revised to reflect the new academic content standards accompanied by setting of achievement levels. "Read aloud" is distinguished from oral administration in South Carolina. Oral administration, which is standardized, is an accommodation permitted for students with disabilities as designated in their IEPs. Oral administration will not be discontinued, as warranted, for these students.

Report:
PACT Grade 3 "Read Aloud" Study, 2006

Vertical Scaling for PACT

From their inception, each of the PACT grades has been scaled separately. Vertically moderated standards were developed, permitting the assumption that achievement levels could be compared across grade levels. For example, the vertically moderated standards allow comparisons of the percentage of students scoring "Basic" in grade 4 and the percentage of students scoring "Basic" in grade 5. With vertically moderated standards, equal weighting of achievement levels across grade levels is justifiable.

Given the requirements of state and federal accountability legislation,

however, there has been interest in vertical calibration and/or scaling of PACT over time. Discussions of vertical calibration of PACT tend to focus on the subjects of English language arts (ELA) and mathematics. These subjects tend to build on similar content from year to year. Science and social studies curricula often vary a great deal from year to year. While sixth-grade ELA may involve a more detailed and sophisticated version of grade five ELA, sixth-grade social studies may be centered on geography, while fifth-grade social studies deals with world history. The common calibration of such diverse subject matter might be theoretically possible, but it would be based more on underlying ability that on common subject matter.

The Office of Assessment has conducted or sponsored several investigations of vertical scaling for PACT ELA and mathematics. The TAC was consulted in 2001 regarding the design of the first linking study. At that time, the TAC advised focusing on adjacent level links and advised that "even without a common scale" predictions could be made. The 2001 TAC proceedings address the equating topic:

> Aside from the technical requirements for equating, several other issues arose that are related to the topic. First among those was the meaning of the scores. For example, a third and an eighth grader with the same mathematics score are very unlikely to share the same underlying mathematical thought patterns, but the score would suggest that they are alike. One panel member suggested calling the process "calibrating" as opposed to "equating" as a way of avoiding the assumption that same scores from different grade levels represent equal capabilities on the part of the students.

In an analysis of data from the spring 2001 PACT administration, Engec, et al. (2001), concluded that a multi-step linking process could be used satisfactorily for ELA and mathematics. Hermann (2005) used similar methodology to analyze data gathered during the spring 2003 PACT administration. This paper provided linking data without offering any evaluative judgment as to how well the process worked. Later, Cohen, et al. (2006) looked at the 2003 data more critically. They concluded that, if PACT tests in ELA and mathematics are vertically calibrated, "Policymakers should

expect fluctuations in growth estimates…. Typical fluctuations may be as small as about 1/6 of one year's average growth, or as high as ½ or even 2/3 of one year's average growth." They also thought that there might be "greater volatility in growth rates at lower levels of aggregation."

Data from these studies were presented to the TAC in 2005. TAC members acknowledged the potential benefits of a vertical scale for some purposes (measuring growth, accountability, and evaluation), but cautioned that it would further complicate the system

> From the schools' perspective, the current rating system, in which a school or district can improve its Absolute Rating and still receive a Below Average Improvement Rating is confusing, and might be improved with a better measure of growth. For accountability purposes, the key questions are how to provide the most accurate information and whether it comes from a vertical scale or from the current categorical method. However, accountability is not the only, or even the primary, issue for schools. A vertical scale would assist them with their own decision-making, research and evaluation projects, which might be enhanced by focusing on achievement growth under various conditions. To reach the state's goals, growth must be accelerated. Perhaps having a better measure of it will help teachers better evaluate and accelerate the growth of their students. Still, the TAC pointed out that a change that solves one problem may create another and that it is important to think through the unintended consequences of a vertical scale. Already the Department must ensure that forms are linked for the same grade from year to year; adding vertical scaling is yet another requirement. Moreover, such scales are potentially misleading and professional development would be needed. For example, with scores on a continuous scale, a high-scoring third grader might appear to score like an eighth grader, but without actually taking any eighth grade items. This can lead to inappropriate conclusions by untrained individuals about what students can and cannot do or where they should be placed for instruction.

TAC discussions focused on some of the technical issues and recognized that

periodic recalibrations must be conducted because the "vertical scale might not be sustained" over time and when content standards change. As an aside, the TAC noted that "North Carolina continues to have a vertical scale, but has discontinued the practice of using it to assess growth for accountability purposes." TAC Proceedings (2005) conclude:

> The TAC is cautious about the use of a vertical scale, recognizing the possibilities for looking at growth or for applications such as those by CRESST and Sanders that rely on a scale. Still, it recommended that if the vertical scale is used, it should be used only for adjacent grades. The discussion also framed vertical scaling as a policy issue: even if a defensible scale has been or can be constructed, that does not mean that it should be used.

After four years of study, The Office of Assessment determined that production of a vertical scale was not advisable for the following reasons:

1. Growth scales would likely be misunderstood and potentially misused, resulting in inappropriate educational decision-making for students.
2. Growth scales would likely lead to greater confusion and would require extensive professional development.
3. Growth scales could not be reported for all subjects so PACT scores would be reported on vertical scales for some subjects and on grade-by-grade scales for others.
4. Vertical scaling would require periodic study, which would result in more testing and a lapse in time for reporting on the vertical scale when studies are conducted.

Volatility estimates from the Cohen, et al. study reinforce this decision; at some grades the error in the estimate of growth is two-thirds of a year of growth.

Reports:

Vertical Equating for the 2003 PACT English Language Arts and Mathematics, 2005

South Carolina Vertical Linking Study, 2005

Conduct a controlled cost and program effectiveness study of the State's readiness for online administration of PACT Mathematics and ELA.

Study the elimination of Constructed Response items on PACT Mathematics and ELA.

These issues are currently undergoing study in the *Study on the Feasibility and Cost of Converting the State Assessment Program to a Computer-Based or Computer-Adaptive Format,* which is expected to be completed in Summer 2007.

Figure 1
Comparison of Grade 4 SC and National NAEP Reading with PACT ELA Percent Proficient or Advanced

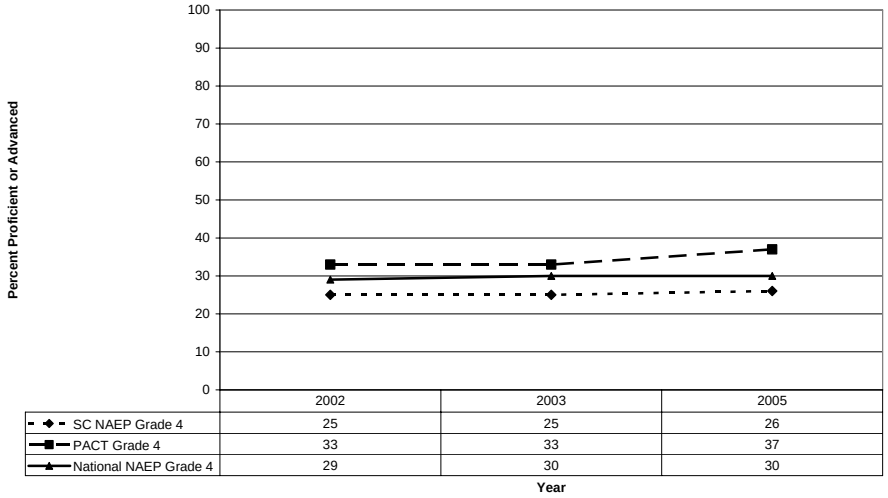


Figure 2
Comparison of Grade 8 SC and National NAEP Reading with PACT ELA Percent Proficient or Advanced

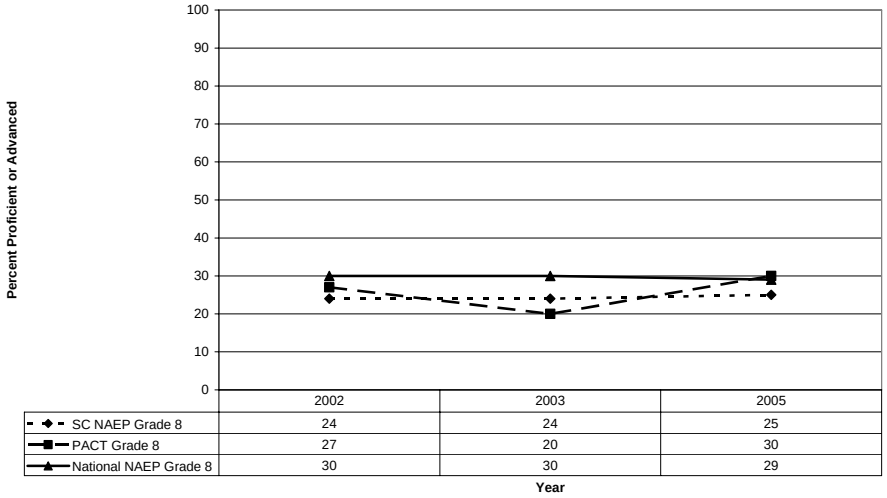


Figure 3
Comparison of Grade 4 SC and National NAEP Reading with PACT ELA Percent Below Basic

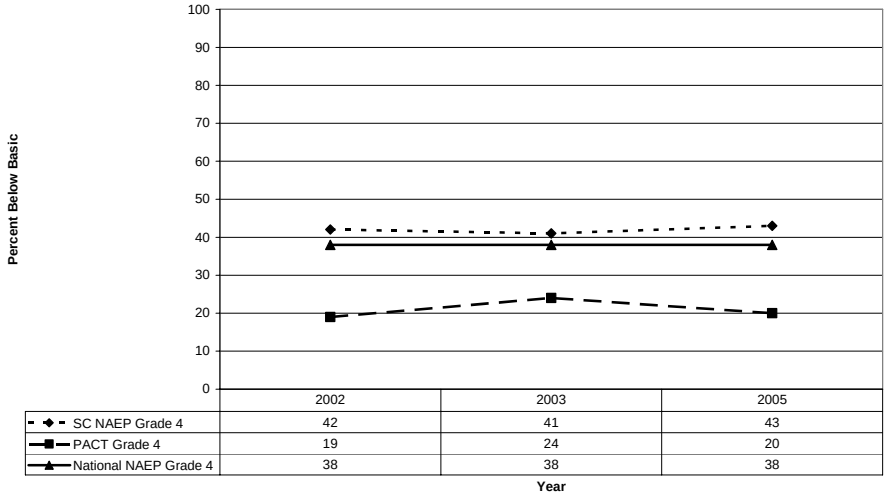


Figure 4
Comparison of Grade 8 SC and National NAEP Reading with PACT ELA Percent Below Basic

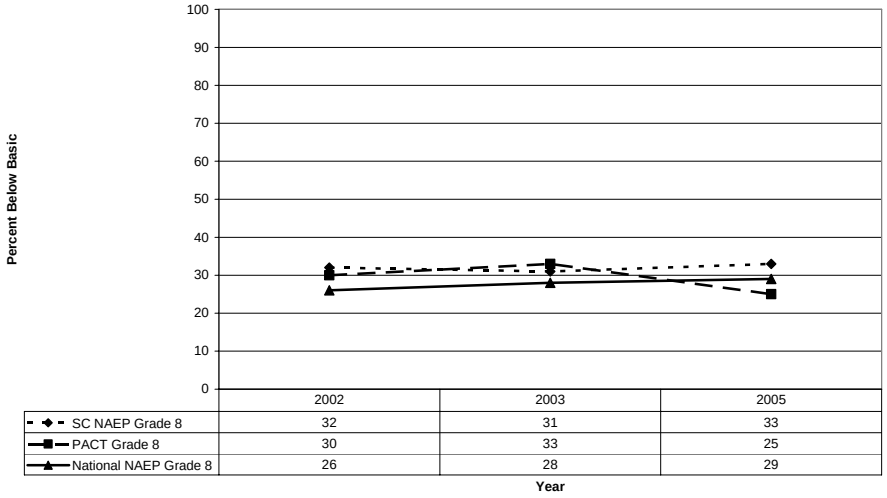
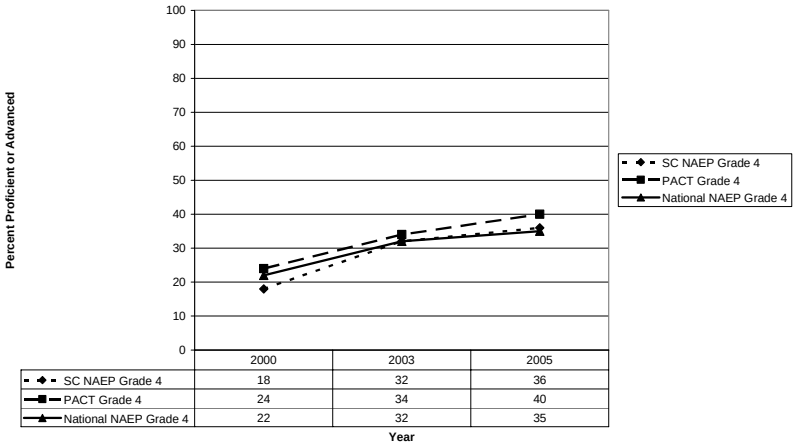


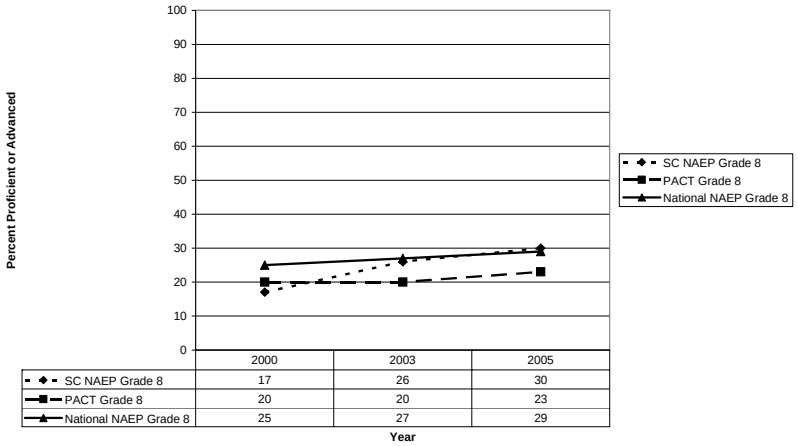
Figure 5
Comparison of Grade 4 SC and National NAEP Math with PACT Math Percent Proficient or Advanced



	2000	2003	2005
SC NAEP Grade 4	18	32	36
PACT Grade 4	24	34	40
National NAEP Grade 4	22	32	35

Year

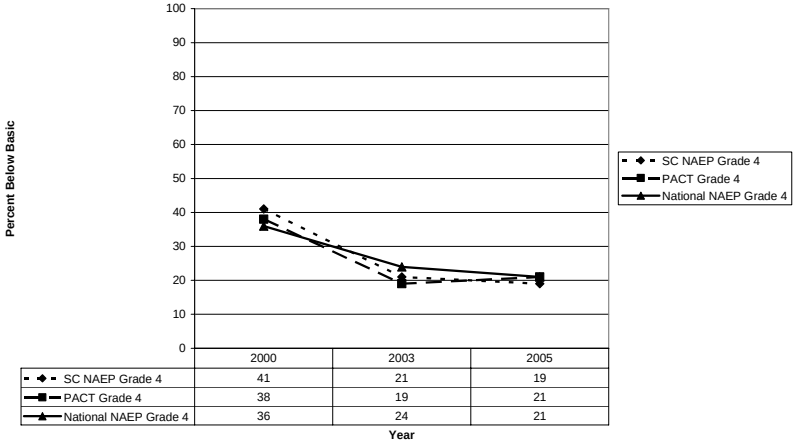
Figure 6
Comparison of Grade 8 SC and National NAEP with PACT Math Percent Proficient or Advanced



	2000	2003	2005
SC NAEP Grade 8	17	26	30
PACT Grade 8	20	20	23
National NAEP Grade 8	25	27	29

Year

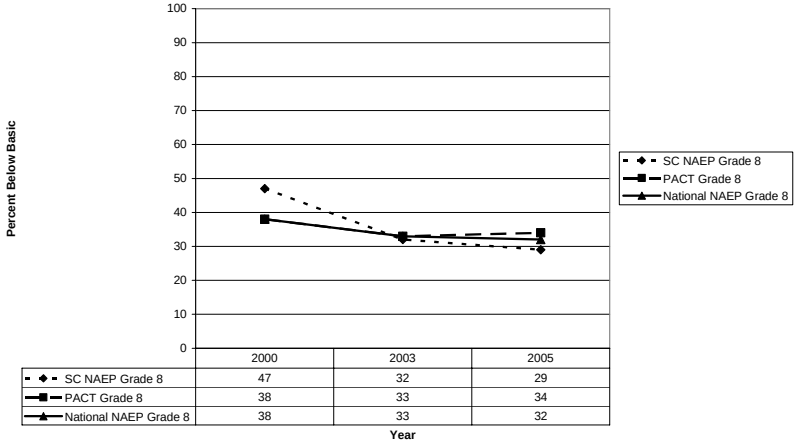
Figure 7
Comparison of Grade 4 SC and National NAEP Math with PACT Math Percent Below Basic



	2000	2003	2005
SC NAEP Grade 4	41	21	19
PACT Grade 4	38	19	21
National NAEP Grade 4	36	24	21

Year

Figure 8
Comparison of Grade 8 SC and National NAEP Math with PACT Math Percent Below Basic



	2000	2003	2005
SC NAEP Grade 8	47	32	29
PACT Grade 8	38	33	34
National NAEP Grade 8	38	33	32

Year

Figure 9
PACT ELA Percent Students Scoring Advanced on Pretest Who Scored Advanced On Posttest - 2000-2005 Longitudinal Data

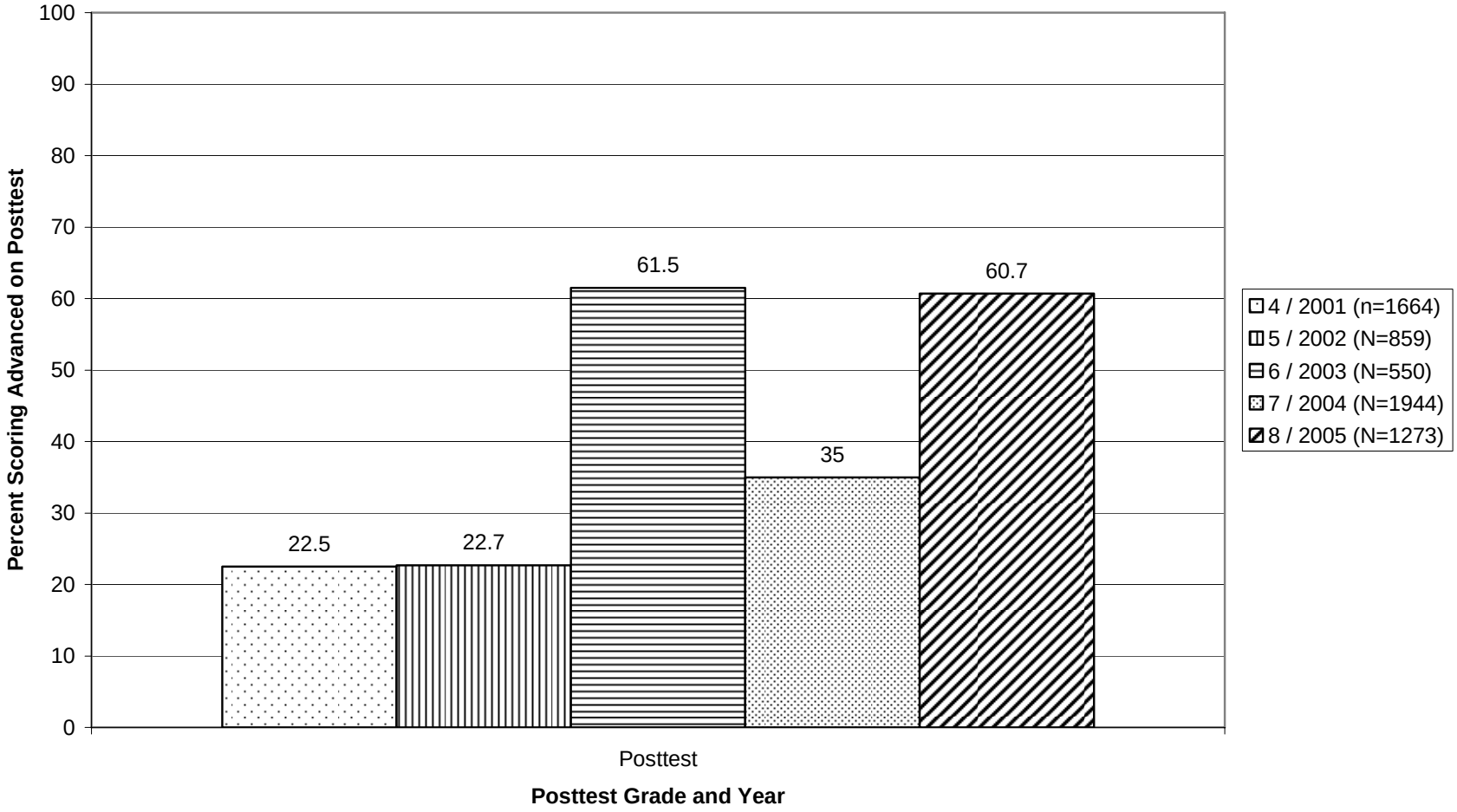
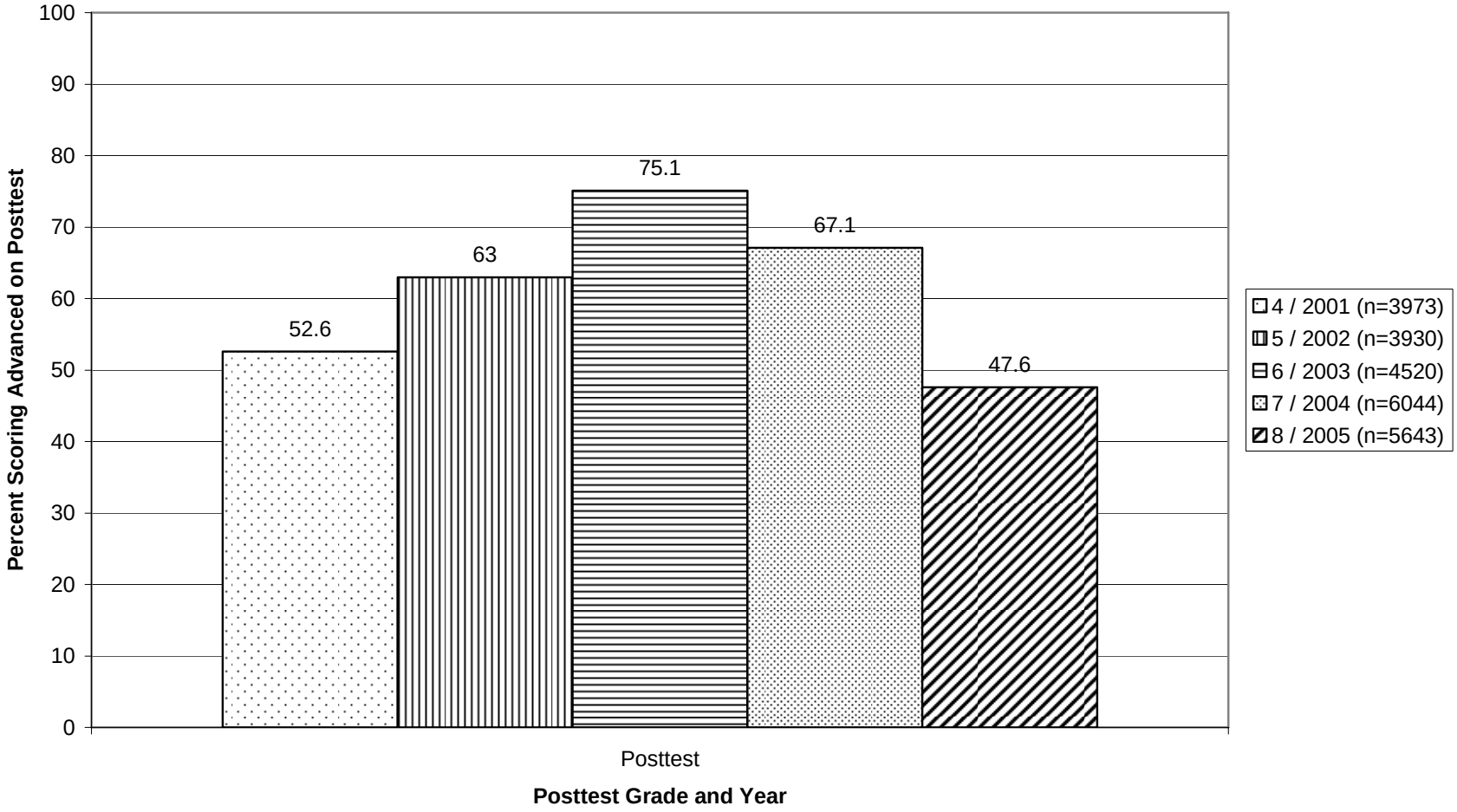


Figure 10
PACT Math Percent Students Scoring Advanced on Pretest Who Scored Advanced on Posttest - 2000-2005 Longitudinal Data



EDUCATION OVERSIGHT COMMITTEE

Subcommittee: Academic Standards and Assessments

Date: September 17, 2007

REPORT/RECOMMENDATION

Review of SC-Alternative ELA and Mathematics Assessments

PURPOSE/AUTHORITY

Section 59-18-320. (A) After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations.
Section 59-18-320 (C). After review and approval by the Education Oversight Committee, the end of course assessments of benchmark courses will be administered to all public school students as they complete each benchmark course.

CRITICAL FACTS

The SC-Alternative ELA and Mathematics field tests were first administered Spring 2006 and revised for the Spring 2007 administration. Recommendations regarding the test following the EOC review must be communicated to the SC State Department of Education, which must respond within one month. State assessments must be reviewed and approved by the Education Oversight Committee.

TIMELINE/REVIEW PROCESS

The SC-Alternative assessments in ELA and Mathematics are intended for administration to students having such severe disabilities that they cannot participate in the PACT or HSAP testing programs. The assessments are administered individually and are designed to assess a broad range of skills expected in the special student population. The SC-Alternative assessment alignment with the academic standards appropriate for students having severe disabilities was assessed by an independent group of experts at the University of North Carolina-Charlotte and at Western Carolina University. The technical aspects of the assessments were evaluated by a measurement expert at the University of South Carolina.

ECONOMIC IMPACT

Cost:

Fund/Source:

ACTION REQUEST

☒ **For approval** ☐ **For information**

ACTION TAKEN

☐ **Approved** ☐ **Amended**

☐ **Not Approved** ☐ **Action deferred (explain)**

2007-2008

SC-ALTERNATIVE ENGLISH LANGUAGE ARTS & MATHEMATICS ASSESSMENTS

From The Division of Accountability

Review of the SC-Alternative English Language Arts and Mathematics Assessments
Executive Summary

This report summarizes the results from studies of the South Carolina Alternate Assessment (SC-Alt) English Language Arts (ELA) and Mathematics field tests administered in Spring 2006 and the revised assessments administered in Spring 2007. The studies were conducted under the auspices of the Education Oversight Committee (EOC) as part of its responsibilities listed in the Education Accountability Act of 1998 (EAA):

> After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations. (Section 59-18-320 A)

The report describes the SC-Alt ELA and Mathematics assessments, describes the studies conducted for this review, presents the findings from the studies, and makes recommendations regarding the assessments.

The SC-Alt ELA and Mathematics assessments are designed for administration to students having significant cognitive disabilities. Students having significant cognitive disabilities function below grade level expectations and have levels of disabilities such that they cannot participate in the regular administrations of the Palmetto Achievement Challenge Tests (PACT) or the High School Assessment Program (HSAP) assessments, even with test accommodations or modifications. Federal No Child Left Behind (NCLB) and Individuals With Disabilities Education Act (IDEA) legislation require that all students be tested and allow for alternative assessment of students having significant cognitive disabilities. The students tested with the SC-Alt ELA and Mathematics assessments represent approximately 0.5% of the total student population in the grade levels tested. The majority of the students to whom the SC-Alt is administered have disabilities classified as Trainable Mental Disability, Profound Mental Disability, or Autism.

The SC-Alt is intended to replace current PACT-Alternate assessments (for grades 3 through 8) and HSAP-Alternate assessment (for grade 10). The SC-Alt assessments are needed to replace PACT-Alt and HSAP-Alt because of changes and clarifications in NCLB regulatory guidance and the reauthorization of IDEA. These changes to federal legislation regarding students with significant cognitive disabilities require that instruction and assessment for these students be based on the same grade level academic standards as apply to all other students, although at less complex levels or with emphasis on prerequisite skills.

The SC-Alt ELA and Mathematics assessments are individually administered to students by teachers during a six- to seven-week window during the Spring of the school year. Each SC-Alt ELA and mathematics test form consists of twelve performance tasks containing four to eight test items each. There are three forms of the test: one for administration to students aged 8 to 10 years (elementary school grades 3 through 8); one for students aged 11 to 13 years (middle school grades 6 through 8), and one for students aged 15 years (high school grade 10). The

test questions are scripted for standardization of administration and administered and scored by the student's teacher; a trained adult monitor unrelated to the student is also present during the test administration.

Two sets of studies were analyzed for the review of the SC-Alt ELA and Mathematics field tests:

- studies of the alignment between the SC-Alt ELA and Mathematics assessments and the state academic standards conducted by University of North Carolina-Charlotte and Western Carolina University professors of curriculum and special education, in cooperation with the South Carolina State Department of Education (SDE) and the National Alternate Assessment Center (Flowers, Browder, Wakeman, & Karvonen, April 2006);
- a technical review of the task and item data from the 2007 test administration conducted by a professor of educational research and assessment at the University of South Carolina.

In addition, EOC staff reviewed and analyzed information and documentation provided by the SDE about the SC-Alt ELA and Mathematics tests.

Conclusions

The studies conducted in this review identified a number of strengths of the SC-Alt ELA and Mathematics assessments:

- ✓ The assessments provide accountability and information for instructional improvement for students having significant cognitive disabilities who would not otherwise be assessed in the state testing programs, even with test accommodations and modifications;
- ✓ With the exception of the ELA Research standard, the assessments are aligned with the same grade level academic standards as for all students, although at levels of complexity appropriate for the diversity of cognitive functioning observed among students having significant cognitive disabilities;
  - o The rationale for not assessing the ELA Research standard which was provided by the SDE and its advisory committee indicated that the Research standard was more appropriately assessed in the course of classroom instruction;
- ✓ The assessments address increasingly complex and more difficult skills across student age levels and have been designed to provide a vertical scale to measure growth;
- ✓ The items in the assessments have a wide range of difficulty and the tests are able to discriminate between high and low levels of ability;
- ✓ The assessments are individually administered by the students' teachers in the familiar context of the classroom;
- ✓ The assessment formats allow students to respond to the items using the communication modes the student uses during instruction, such as oral response, pointing, use of eye gaze, use of a response card, sign language, or an augmentative communication device;
- ✓ The assessments are scripted, their administration and scoring is observed by monitors, and the teachers and monitors administering the assessments undergo training to ensure that the assessment administration is standardized and the results are valid measures of the student's ability;
- ✓ The assessments are administered over a six- to seven-week period, providing flexibility and opportunities for maintaining student motivation and interest and reducing student fatigue;

- ✓ The procedures for placing the student at the appropriate level for beginning each assessment reduces student fatigue and maximizes students' opportunities to show their highest performance.

Some concerns were also identified through this review:

- ✓ The analysis of the technical quality of the assessments revealed that approximately one-third of the items were "flagged" for having statistical values outside the expected range, although most of the flags were for relatively minor statistical differences;
- ✓ However, approximately 15 items were flagged for Differential Item Functioning, a measure which suggests that an item's wording or content may confer an advantage to one subgroup of test-takers compared to another subgroup;
- ✓ The authors of the alignment study indicated that a draft teacher's guide to the alternate assessments provided for the alignment study was out of date and needed to be updated to address changes to the academic standards and the alternate assessments.

Recommendations

Overall, the SC-Alt ELA and Mathematics assessments are aligned with the South Carolina ELA and Mathematics academic standards and have acceptable technical quality consistent with the requirements of Section 59-18-320 A. Based on these findings, it is recommended that the SC-Alternate ELA and Mathematics assessments be approved with the following recommendations:

1. The South Carolina State Department of Education (SDE) should review the SC-Alt ELA and Mathematics items which were "flagged" for their statistical values, especially those items flagged for Differential Item Functioning, to identify reasons for the statistical aberrations observed and to identify the need to revise or eliminate the items from the assessments.
2. The SDE should develop and disseminate updated professional development guides and materials related to the Assessment Standards and Measurement Guidelines and the SC-Alt assessments, including information to assist teachers to align their instruction with the Assessment Standards and Measurement Guidelines.

Introduction

This report summarizes the results from studies of the South Carolina Alternate Assessment (SC-Alt) English Language Arts (ELA) and Mathematics field tests administered in Spring 2006 and the revised assessments administered in Spring 2007. The studies were conducted under the auspices of the Education Oversight Committee (EOC) as part of its responsibilities listed in the Education Accountability Act of 1998 (EAA):

> After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee, established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations. (Section 59-18-320 A)

The report describes the SC-Alt ELA and Mathematics assessments, describes the studies conducted for this review, presents the findings from the studies, and makes recommendations regarding the assessments.

Development of SC-Alt ELA and Mathematics Assessments

The SC-Alt ELA and Mathematics assessments are intended for administration to students having significant cognitive disabilities. These students, who are functioning below grade level expectations, have levels of disabilities such that they cannot participate in the regular administrations of the Palmetto Achievement Challenge Tests (PACT) or the High School Assessment Program (HSAP) assessments, even with accommodations or modifications. Federal No Child Left Behind (NCLB) and Individuals With Disabilities Education Act (IDEA) legislation require that all students be tested and allow for alternative assessment of students having significant cognitive disabilities.

The SC-Alt is intended to replace current PACT-Alternate assessments (for grades 3 through 8) and HSAP-Alternate assessment (for grade 10). The SC-Alt assessments are needed to replace PACT-Alt and HSAP-Alt because of changes and clarifications in NCLB regulatory guidance and the reauthorization of IDEA. These changes to federal legislation regarding students with significant cognitive disabilities require that instruction and assessment for these students be based on the same grade level academic standards as apply to all other students, although at less complex levels or with emphasis on prerequisite skills.

The current PACT-Alt and HSAP-Alt assessments are based on alternative academic standards rather than the grade level academic standards. In practice, this has meant that teachers have based instruction and assessment largely on the academic standards for grades Kindergarten through grade 2 regardless of the age levels of their students having significant cognitive disabilities. The federal changes have also led to changes in goals for Individualized Education Programs (IEPs) for students with disabilities from individual objectives to objectives based on the state academic standards. To meet federal requirements, the assessments for students having significant cognitive disabilities needed to be revised, and SC-Alt has resulted from those revisions.

Alternative assessments such as SC-Alt are based on state grade level academic standards, but at lower levels of complexity or with greater focus on introductory or prerequisite skills. In 2005 committees composed of ELA and mathematics content specialists, experts in the instruction of significantly cognitively disabled students, and staff from the South Carolina State Department of Education (SDE) and its testing contractor, the American Institutes for Research (AIR), reviewed the academic standards to identify the "standards they felt based on professional judgment were the most important to the population now and in the future" (Overview of the SC-Alt Technical Documentation Presented to the National Alternate Assessment Center, March 16, 2007, p. 6). Following their identification of the priority standards for students having significant cognitive disabilities, these committees developed Assessment Standards and Measurement Guidelines (ASMG) in each subject area to guide instruction and the construction of SC-Alt (the ASMGs are available at http://ed.sc.gov/agency/offices/assessment/programs/SWD/SC-AltAssessmentStandardsandMeasurementGuidelines.html). The SC-Alt ELA and Mathematics assessments are based on the corresponding ASMG, providing a link from the assessment to the state grade level academic standards.

Description of the SC-Alt ELA and Mathematics Assessments

The SC-Alt is individually administered to each student, generally by the teacher who has provided instruction to that student. In addition to the teacher administering the assessment, a trained monitor unrelated to the student must be present during the test administration. The monitor is required to ensure that the assessment is administered and scored properly. The assessment is administered during a 6-7 week window starting in March. The student may complete the assessment for each subject area in one session or, if the student tires or is non-attentive, the assessment may be administered over several days.

The SC-Alt ELA and Mathematics assessments are designed for administration to three age ranges of students rather than to students classified at specific grade levels. An "elementary" form is intended for use with students aged 8-10 by September 1 of the school year of testing (corresponding to the grade band 3 through 5). The "middle school" form is administered to students aged 11-13, corresponding to grade band 6 through 8, and the "high school" form is administered to students aged 15 (the age when most students are classified as 10th graders). The SC-Alt is designed to provide a continuous scale of increasing difficulty for students aged 8 through 13 and age 15, with the content of the test appropriate for students aged 8 through 15. This design is intended to provide appropriate age-related content to maintain interest and motivation on the part of the student being tested.

Each grade-band form consists of 12 performance tasks, with each task containing 4 to 8 items. The performance task format was chosen for the SC-Alt based on the advice of special education advisory committees, based in part on educators' acceptance of the current performance task format of the HSAP-Alt. The PACT-Alt was based on the collection and scoring of a portfolio of student work or behavior. The portfolio format was criticized by educators because of paperwork loads and concerns about the subjectivity of portfolios and their scoring.

The SC-Alt assessment is scripted, with specific directions to the teacher for administration and scoring of the assessment (see Figure 1 for descriptive information about the SC-Alt tasks and items).

Figure 1
SC-Alt Tasks and Items

A task is a set of four to eight related activities, called items. The responses to the items provide evidence of what students know and can do.

- ✓ Each task begins with an introductory statement that establishes the context for what the student will be doing. There is a clear progression within each task from one activity to another.
- ✓ The teacher uses scripted directions to pose specifically worded questions and prompts to the student.
- ✓ The student responds by using the mode of communication that he or she uses during instruction. These response modes include but are not limited to an oral response, pointing, use of eye gaze, a response card, sign language, or an augmentative communication device.
- ✓ The test administrator will use various materials to administer a task or an item to help a student respond. Some of the materials are provided with each task, and some materials that are readily available at the school are provide by the test administrator.
- ✓ The materials may include poster, charts, tables, schedules, and signs that the administrator reads aloud and manipulatives such as checkers, balls, and geometric shapes.
- ✓ Unless the task is presented entirely through the use of concrete objects, resources will also include a set of response cards for each item to facilitate a student's response.
- ✓ Each task addresses one or more of the assessment standards or measurement guidelines.
- ✓ The SC-Alt assesses selected standards or measurement guidelines. Individual students are assessed on a sample of standards and guidelines.

Scripted items:

- ✓ Each item begins with a scripted opening statement in Say/Do format. For example, "Say: Here is a …," or, "Say: Look at/touch the …"
- ✓ The opening statement is followed by a directive for the student to tell or show the teacher which one of several response options is correct. For example, "Say: Tell (show) me what the boy in the story did when he got home."

(Sources: Spring 2006 and Spring 2007 Test Administration Manuals.)

The tasks are ordered in difficulty, with the least complex task appropriate for the student administered first, and, as the student successfully answers the items in each successive task, the testing session is continued through the more complex tasks until the student fails to correctly answer or respond to a specified number of items. Prior to the administration of the SC-Alt for each content area, each student's ability in that content area is evaluated by the teacher using the Student Placement Questionnaire (SPQ) (Appendix 1) to determine the student's entry into the test form (e.g., the first task which will be administered to the student). The teacher's evaluation of the student on the SPQ instrument is based on the teacher's experience during the year of instruction he or she has provided the student. Based on the teacher's evaluation of the student's ability using the SPQ, the student may start the test with the first task, or, if the student has higher levels of cognitive functioning, at task 3 or task 6, as appropriate. This adaptation of the test to the student's abilities is intended to increase the accuracy of the student's test score by only administering appropriately challenging items to the student. The use of the SPQ is also intended to avoid excessively tiring the student and to

maintain the student's interest and motivation by avoiding items that are well below the student's ability level. If the teacher finds that the beginning task suggested by the SPQ is too challenging for the student, the teacher chooses a lower level task based on the criteria listed in the administration directions. Regardless of the student's entry point into the assessment, each student must complete at least 5 tasks, but may respond to more than 5 tasks if the student's performance meets the criteria for continuing.

The student's response to each question on the assessment is recorded and scored by the teacher administering the assessment. The test administrators and monitors must receive professional development on the administration and scoring of the assessment. The scoring of each item may be "scaffolded" if the student provides an incorrect answer or does not respond. For example, if an item has three answer options, only one of which is correct, and the student fails to choose the correct answer on the first try, on the student's second try the teacher may restate the question but provide only two responses, eliminating the incorrect answer chosen initially by the student. If the student again fails to choose the correct answer (or does not respond to the question), then the teacher records a "0" or "No Response" and moves on to the next item. If the student correctly responds when only two choices are given rather than three choices, the student is awarded fewer points than if he or she had correctly answered the item on the first try. This scaffolding of the scoring provides for a level of success for the student and allows the identification of the student's partial level of skill or knowledge in the standard assessed by the item.

Studies Conducted of SC-Alt ELA and Mathematics Assessments

The SC-Alt ELA and Mathematics assessments were initially field tested in Spring 2006. The tasks and items in the initial field test were selected for further use, revised, or eliminated following reviews by content area committees, reviews of data from the technical analyses of the task and item data, reviews of the results of the study of the task and item alignment with the academic standards, and reviews of comments from teachers who had administered the field tests. Following this review, three grade-band forms (grades 3-5, grades 6-8, and grade 10) for each content area were created using the revised tasks and items from the 2006 field test for administration in Spring 2007. The studies conducted for this review are based on data from the 2006 field test and from the 2007 administration of the revised tasks and items.

Studies of the alignment between the SC-Alt ELA and Mathematics assessments and the state academic standards were conducted by University of North Carolina-Charlotte and Western Carolina University professors of curriculum and special education, in cooperation with the SDE and the National Alternate Assessment Center (Flowers, Browder, Wakeman, & Karvonen, April 2006). The studies were part of a project to develop and pilot alignment procedures designed for evaluating tests for students having significant cognitive disabilities. The alignment studies were conducted in Spring 2006.

A technical review of the task and item data from the 2007 test administration was conducted by a professor of educational research and assessment at the University of South Carolina. In addition, EOC staff reviewed and analyzed information and documentation provided by the SDE about the SC-Alt ELA and Mathematics tests (the documentation provided is listed in Appendix 2).

Findings

Numbers of Students Assessed and Numbers of Tasks and Items Administered

The numbers and the disability classifications of students participating in the 2006 field test and in the 2007 administration of SC-Alt ELA and Mathematics assessments are listed in Table 1. The eligibility of students to participate in the SC-Alt assessments is based upon meeting the criteria listed in Appendix 3. Students eligible to participate in the SC-Alt assessments have significant cognitive disabilities and represent approximately 0.5% of all students enrolled in grades 3 through 8 and grade 10, and approximately 4% of all special education students.

Table 1
Numbers of Students Tested and Their Disabilities, 2006 Field Test and 2007 Administration of SC-Alt ELA and Mathematics Assessments

Disability Classification	Number Students Participating in 2006 Field Test (%)	Number Students Participating in 2007 Administration (%)
Trainable Mentally Disabled (TMD)	973 (51.2)	992 (40.1)
Autism	277 (14.6)	406 (16.4)
Profound Mentally Disabled (PMD)	265 (13.9)	273 (11.0)
Educable Mentally Disabled (EMD)	194 (10.2)	546 (22.1)
Other*	191 (10.0)	259 (10.5)
Totals	1,900 (100)	2,476 (100)

Note: Totals may not equal 100% due to rounding.
* Includes categories: Multiple Disability; Other Health Impaired; Traumatic Brain Injury; Hearing, Visual, Speech, or Language Disabled; Orthopedically Impaired; Learning Disability; Unknown.

Many of the tasks and items administered in the Spring 2006 field test were revised or eliminated based on the academic standard alignment studies and the review of the technical characteristics of the items, so the data from the Spring 2007 administration of the SC-Alt ELA and Mathematics assessments were used for the technical analysis of the assessment items in this review. The numbers of tasks and items administered in Spring 2007 and reviewed in this report are listed in Table 2.

Table 2
Numbers of Tasks and Items By Grade Band Form
SC-Alt ELA and Mathematics 2007 Administration

Content Area	Grade Band 3-5 Form		Grade Band 6-8 Form		Grade 10 Form		Total No. Tasks	Total No. Items
	No. of Tasks	No. of Items	No. of Tasks	No. of Items	No. of Tasks	No. of Items		
ELA	12	68	12	65	12	64	36	197
Mathematics	12	53	12	55	12	60	36	168

Study of the Alignment of the SC-Alt Items to the State Academic Standards

During the spring of 2006 the SC-Alt ELA and Mathematics field test tasks and items were reviewed by a group of experts at the University of North Carolina-Charlotte and at Western Carolina University in partnership with the National Alternative Assessment Center (Fowler, et al., 2006). The Executive Summary from the alignment study report is provided in Appendix 4. The purpose of the review was to evaluate the alignment of the assessment items with the state academic standards using a pilot set of criteria for evaluating the alignment of assessments intended for use with students having significant cognitive disabilities. The review results were also used by the SDE and its contractor, the American Institute for Research (AIR) in the evaluation of the field test items for future use on the operational forms of SC-Alt.

Seven alignment criteria were developed by a team of content experts, special educators, and measurement experts. The alignment criteria were similar to other criteria for evaluating the alignment of test items to academic standards, but included three additional criteria (criteria 5-7) designed to apply to assessments intended for students having significant cognitive disabilities. The alignment criteria used in the study are listed in Table 3.

Table 3
Criteria for Judging the Alignment of Assessment Items and Academic Standards

1. The content is academic and includes the major domains/ strands of the content area as reflected in state and national standards (e.g., reading, math, science.)
2. The content is referenced to the student's assigned grade level (based on chronological age).
3. The achievement expectation is linked to the grade level content, but differs in depth or complexity; it is not grade level achievement. It may focus on prerequisite skills or those learned at earlier grades, but with applications to the grade level content. When applied to state level alternate assessments, these priorities are accessible to IEP planning teams.
4. There is some differentiation in achievement across grade levels or grade bands.
5. The focus of achievement promotes access to the activities, materials, and settings typical of the grade level but with the accommodations, adaptations, and supports needed for individualization.
6. The focus of achievement maintains fidelity with the content of the original grade level standards (content centrality) and when possible, the specified performance (category of knowledge).
7. Multiple levels of access to the general curriculum are planned so that students with different levels of symbolic communication can demonstrate learning. (Fowler, et al., 2006, p. 11)

Using these seven criteria, a team composed of three English language arts experts, two mathematics experts, two experts in the education of students having significant cognitive disabilities, and two experts in educational measurement evaluated the 44 ELA tasks consisting of 248 items and the 44 mathematics tasks containing 216 items used in the Spring 2006 SC-Alt field test. These tasks and items provided the basis for the creation of 2007 forms for grade bands 3 through 5, 6 through 8, and grade 10. The evaluators also administered a Curriculum

Indicators Survey to a small sample of South Carolina teachers of significantly disabled students to assess classroom instruction.

Following training in the seven alignment criteria, the evaluators achieved approximately 80% exact inter-rater agreement for the ELA items and greater than 80% agreement for the mathematics items, suggesting that the criteria were clear and that the alignment evaluations provided through the process were reliable.

With regard to criteria 1 and 2, all but 16 of the ELA items were found to be assessing academic skills; these 16 items were eliminated from further consideration, leaving 232 ELA items in the study. All of the mathematics items were found to be academic. Twelve of the 16 ELA items judged to be non-academic were deleted from the item pool and not used for the creation of operational forms of the tests. Most of the remaining items judged non-academic were the first items administered at the beginning of the least complex tasks and served either to introduce the topic of the task or to identify the student's engagement in the assessment activity. These "engagement" items were scored by the teacher using a scoring rubric having values from "Student demonstrates sustained involvement in the activity…" (indicating the maximum level of engagement) to "Student does not demonstrate any awareness of the object(s) or involvement in the activity taking place or may refuse to engage in the activity at any level" (non-responsive, or the lowest level of engagement).

The ELA items were judged to be aligned to both the National Council of Teachers of English (NCTE) standards and the South Carolina ELA academic standards, with the exception of Research. None of the ELA items addressed the Research standard. In response to this finding, the SDE and its alternate assessment advisory committee provided the following rationale for the exclusion of Research from the SC-Alt ELA assessment:

> "Rationale for Omission of the Research Domain from the SC-Alt Assessment Standards and Measurement Guidelines (ASMGs)
>
> A committee of special educators and English language arts (ELA) content specialists reviewed the state academic grade level standards to determine those that were most appropriate for students with significant cognitive disabilities both now and in the future. This group was committed to ensuring that the SC-Alt addressed the depth and breadth of the academic standards and selected standards for inclusion in the Assessment Standards Measurement Guidelines document with this in mind.
>
> The committee determined that although the Communication Goal is not directly assessed in the general assessment, this area is critical for students who do not transfer skills readily and must be taught communication skills in a variety of ways and settings. Therefore, they recommended that the assessment include tasks from the Communication Goal. The committee examined the Research Goal and standards carefully and determined that although these are important for daily classroom instruction and assessment, they are difficult to assess in a performance task. The committee was comfortable with recommending that the assessment standards focus primarily on the Reading, Writing, and Communication goals for the ELA portion of the SC-Alternate Assessment." (SDE, no date)

Most of the ELA items assessed Reading (approximately 80%), followed by Writing (13%-20% depending on the grade band form) and Communication (3%-10% depending on the grade band).

All of the mathematics items addressed the grade band standards. The Number and Operations standard was most frequently assessed (31%-34% of the items depending on the grade band form), followed by Measurement (20%-29% of the items), Geometry (20%-21% of the items), Algebra (11%-14% of the items) and Data Analysis and Probability (9%-12% of the items).

With regard to alignment criterion 3, the evaluators found that the Assessment Standards and Measurement Guidelines and the test items for both ELA and Mathematics satisfied the criterion that the assessment be linked to grade level standards but at a lower level of complexity. However, the evaluators expressed some concern about the emphasis on Reading in the ELA items and the emphasis on Number and Operations among the mathematics items, along with the relatively low levels of cognitive demand presented by the items and the low levels of cognitive expectations for students during instruction which was reported by the teachers. The evaluators indicated that this issue should be discussed to determine the need for broadening the curriculum for the students. The evaluators also identified the need to provide professional development to teachers on how to increase the cognitive complexity of instructional activities.

The evaluators found that there is significant differentiation across the grade bands in the complexity of achievement measured by the ELA and mathematics items (criterion 4). With regard to criterion 5, the evaluators found that the ELA and mathematics tasks and items were appropriate for the target group of students and that the items, as intended, were appropriate for either younger or older students. The evaluators did note, however, that the existing professional development materials were based on standards and instructional strategies and materials from the Kindergarten to second grade standards. The evaluators recommended that revised professional development materials be developed to assist teachers to adapt grade level activities to their students' cognitive capacities and skill levels.

The ELA and mathematics tasks and items were found to be well aligned with the content and cognitive skills found in the grade level academic standards (criterion 6). The evaluators also recommended that professional development materials designed to help teachers identify the alignment of their instructional objectives and the state academic standards be created and disseminated.

Finally, with regard to criterion 7, that the tasks and items address the full range of student communication skills, the evaluators found that construction of the ELA and mathematics items was weighted heavily toward students who possess a higher level of communication skill (i.e., at the symbolic level). The evaluators identified four levels of communication skills among students having significant cognitive disabilities:

1. Awareness: student has no clear response and no objective in communication.
2. Pre-symbolic: student communicates with gestures, eye gaze, purposeful moving to object, sounds.
3. Early Symbolic: student begins to use pictures or other symbols (less than 10) to communicate within a limited vocabulary.
4. Symbolic: student speaks or has vocabulary of signs, pictures to communicate. Recognizes some sight words, numbers, etc. (Fowler, et al., 2006, p. 37)

The evaluators questioned whether the assessments could identify the proficiency of students communicating at lower levels than the symbolic level when the tasks and items were weighted so heavily toward symbolic communication.

Overall, the evaluators judged that the assessment system "links to the grade level content" (Fowler, et al., 2006, p. 7) and that the evidence from the assessments supports the judgment that the Assessment Standards and Measurement Guidelines and the ELA and mathematics tasks and items meet all seven alignment criteria. The evaluators recommended that the professional development materials provided by the SDE at the time of the evaluation study be revised to reflect the current focus of the Assessment Standards and Measurement Guidelines and the SC-Alt assessments.

Technical Analysis of Test Forms, Tasks, and Items

Dr. Christine DiStefano, a professor of educational research and measurement at the University of South Carolina, conducted a review of the technical characteristics of the SC-Alt ELA and Mathematics assessments. Dr. DiStefano's studies focused on the evidence provided from the technical data which informed the requirement in the Education Accountability Act (Section 59-18-320A) that the assessments be reviewed for their "level of difficulty and validity" and "the ability to differentiate levels of achievement." Her report is included in Appendix 5 of this report.

Dr. DiStefano stated that a strength of the SC-Alt was the use of multiple measures both to identify students for administration of the SC-Alt (the student participation guidelines) and to determine the starting point among the assessment tasks for individual students (the Student Placement Questionnaire). She also noted that the training provided for test administrators on placement of students on the test and scoring of their responses helped to ensure the validity of the test scores.

Dr. DiStefano found that the ELA and mathematics item statistics were within acceptable ranges for the intended use of the tests. As intended, the tests increased in difficulty across the grade bands, indicating that older students were assessed on more complex skills than younger students. Overall, the assessments were of moderate difficulty, with students answering approximately 60% of the items correctly. The item statistics indicated that the tests had acceptable levels of discrimination, indicating that both the ELA and mathematics assessments provided results which were useful to distinguish between high and low ability students.

The technical analysis revealed that approximately one-third of the test items were "flagged" for having technical statistics which exceeded the expected ranges. Most of the "flags" were considered to be for rather minor departures from the technical expectations, but at least 15 items showed Differential Item Functioning (DIF) statistics possibly indicating that some characteristics of the items enabled one demographic group to score higher on the items than another demographic group. Dr. DiStefano indicated that this potential "bias" of the item toward one group in favor of another should be investigated by reviewing the item statistics and the wording and content of the items to identify potential reasons for the DIF flag. All of the items chosen for the test forms were reviewed and approved by a "bias review committee," but the empirical DIF statistics suggest there may some unanticipated explanation for the differential performance of subgroups. Dr. DiStefano also pointed out that the item statistics may have been affected by the small sample sizes, especially with the grade 10 form; smaller sample sizes for calculating the statistics increase the size of the margins of error in estimating the true values of the statistics.

Finally, Dr. DiStefano recommended that the outcomes from the SC-Alt ELA and Mathematics assessments be reviewed when impact data are available to evaluate the overall difficulty of the operational assessments and the rigor of the performance standards. Based on the data

available at this time, however, she found that the SC-Alt appears to perform effectively to assess South Carolina's students with significant cognitive disabilities.

Conclusions and Recommendations

The studies conducted in this review identified a number of strengths of the SC-Alt ELA and Mathematics assessments:

- ✓ The assessments provide accountability and information for instructional improvement for students having significant cognitive disabilities who would not otherwise be assessed in the state testing programs, even with test accommodations and modifications;
- ✓ With the exception of the ELA Research standard, the assessments are aligned with the same grade level academic standards as for all students, although at levels of complexity appropriate for the diversity of cognitive functioning observed among students having significant cognitive disabilities;
  - o The rationale for not assessing the ELA Research standard which was provided by the SDE and its advisory committee indicated that the Research standard was more appropriately assessed in the course of classroom instruction;
- ✓ The assessments address increasingly complex and more difficult skills across student age levels and have been designed to provide a vertical scale to measure growth;
- ✓ The items in the assessments have a wide range of difficulty and the tests are able to discriminate between high and low levels of performance;
- ✓ The assessments are individually administered by the students' teachers in the familiar context of the classroom;
- ✓ The assessment formats allow students to respond to the items using the communication modes the student uses during instruction, such as oral response, pointing, use of eye gaze, use of a response card, sign language, or an augmentative communication device;
- ✓ The assessments are scripted, their administration and scoring is observed by monitors, and the teachers and monitors administering the assessments undergo training to ensure that the assessment administration is standardized and the results are valid measures of the student's ability;
- ✓ The assessments are administered over a six- to seven-week period, providing flexibility and opportunities for maintaining student motivation and interest and reducing student fatigue;
- ✓ The procedures for placing the student at the appropriate level for beginning each assessment reduces student fatigue and maximizes students' opportunities to show their highest performance.

Some concerns were also identified through this review:

- ✓ The analysis of the technical quality of the assessments revealed that approximately one-third of the items were "flagged" for having statistical values outside the expected range, although most of the flags were for relatively minor statistical differences;
- ✓ However, approximately 15 items were flagged for Differential Item Functioning, a measure which suggests that an item's wording or content may confer an advantage to one subgroup of test-takers compared to another subgroup;
- ✓ The authors of the alignment study indicated that a draft teacher's guide to the alternate assessments provided for the alignment study was out of date and needed to be updated to address changes to the academic standards and the alternate assessments.

Recommendations

Overall, the SC-Alt ELA and Mathematics assessments are aligned with the South Carolina ELA and Mathematics academic standards and have acceptable technical quality consistent with the requirements of Section 59-18-320 A. Based on these findings, it is recommended that the SC-Alternate ELA and Mathematics assessments be approved with the following recommendations:

1. The South Carolina State Department of Education (SDE) should review the SC-Alt ELA and Mathematics items which were "flagged" for their statistical values, especially those items flagged for Differential Item Functioning, to identify reasons for the statistical aberrations observed and to identify the need to revise or eliminate the items from the assessments.
2. The SDE should develop and disseminate updated professional development guides and materials related to the Assessment Standards and Measurement Guidelines and the SC-Alt assessments, including information to assist teachers to align their instruction with the Assessment Standards and Measurement Guidelines.

APPENDIX 1

Example of Student Placement Questionnaire

SC-ALT STUDENT PLACEMENT QUESTIONNAIRE
ENGLISH LANGUAGE ARTS

(completed SPQ example)

Follow steps 1-4 to complete the SPQ and identify the starting task.

(1) Please darken the bubble (●) that corresponds to the most appropriate response for this student. Mark only one response for each item. Please mark a response for all items below. Use a no. 2 pencil only.

	No, she/he cannot do this	With physical prompting/hand-over-hand	With verbal/gestural prompting	Independently
In reading, can this student:				
1. Attend to text read aloud?	○	●	○	○
2. Recall details in text read aloud?	○	●	○	○
3. Recognize some high-frequency written words?	○	○	●	○
4. Draw conclusions or make inferences about texts?	○	○	○	●
In writing, can this student:				
5. Write his or her name using a pencil, name stamp, letter tiles, or other means?	○	●	○	○
6. Use objects, pictures, and/or picture symbols to write in any format?	○	○	●	○
7. Copy, trace, or print letters?	○	●	○	○
8. Use oral language and/or letters and words to write?	○	○	○	●
In communicating, can this student:				
9. Listen (i.e., demonstrate receptive behavior) and respond?	●	○	○	○
10. Participate in conversations by responding appropriately?	○	●	○	○
11. Use language to express a preference, opinion, or viewpoint?	○	●	○	○
12. Recognize and understand the meaning of environmental signs (e.g., street signs, store signs, school signs)?	○	●	○	○
(2) Write in the total number of bubbles you marked in each column		1	7	2
		col. 1	col. 2	col. 3

(3) Calculate the SPQ total score

(a) write the column totals from (2) in (a) below
(b) multiply and write the results in (b) below
(c) sum the results from (b) and write the sum in (c)

	(a)		(b)
Column 1 Total	1	x 3 =	3
		+	
Column 2 Total	7	x 2 =	14
		+	
Column 3 Total	2	x 1 =	2
(c) Total SPQ Score =			0 9

(4) Identify the starting task for this student using the SPQ total score from step (3).

If the total score is in this range	Start at this task	Administer all items in at least these tasks
0-11	Task 1	1-5
12-22	Task 3	3-9
23-36	Task 6	6-12

APPENDIX 2
SC-Alt Documentation Provided by
South Carolina Department of Education

1. American Institutes for Research. *Plan for Setting Status Based Performance Standards for SC-Alt* (April 2007)
2. American Institutes for Research. *SC-Alt ELA and Math Operational Assessments Proposed Grade Band Design for Operational Administration and Linking* (August 9, 2006)
3. American Institutes for Research. *South Carolina Alternate Assessment: Marginal Reliability Estimates & Standard Error of Measurement Across Grade Bands and Content Areas* (August 14, 2007)
4. American Institutes for Research. *Technical Report (Draft), SC-Alt Setting Standards in Grade Bands 3-5, 6-8 and 10, Spring 2007 Standard Setting* (June 2007)
5. *Assessment Standards and Measurement Guidelines, SC-Alt English Language Arts*, March 2006
6. *Assessment Standards and Measurement Guidelines, SC-Alt Mathematics,* March 2006
7. *Descriptions of Achievement Levels, ELA, Mathematics, Science, Social Studies* (no date)
8. *District Test Coordinator's Supplement, SC-Alt English Language Arts, Mathematics,* Spring 2007
9. Flowers, C., Browder, D., Wakeman, S, & Karvonen, M. (April 2006). *Alternate Assessment Alignment Pilot Study Report to the South Carolina State Department of Education.* University of North Carolina at Charlotte
10. *Information and Procedures for the SC-Alt Field Test Item Data Review Meeting* (May 19, 2006)
11. *Notes from Alignment Study RE: ELA Tests* (no date)
12. *Overview of SC-Alt Technical Documentation Presented to the National Alternate Assessment Center* (March 16, 2007)
13. *Rationale for Omission of the Research Domain from the SC-Alt Assessment Standards and Measurement Guidelines* (no date)
14. *South Carolina Alternate Assessment 2006 Field Test Brief for the August Meeting of the South Carolina Department of Education Technical Advisory Committee* (August 7, 2006)
15. *Summary of the Data Review Comments, Outcomes from the Item Analysis, Teacher Comments, and Alignment Study* (no date)
16. *Summary Tables for SC-Alt Technical Advisory Committee, South Carolina Alternate Assessment, 2007 Administration* (July 2007)
17. *Synopsis for State Board of Education (Draft)* (September 12, 2007)
18. *Test Administration Manual, SC-Alt Field Test, Spring 2006*
19. *Test Administration Manual, SC-Alt English Language Arts, Mathematics, Spring 2007*
20. *The South Carolina Alternate Assessment Spring 2007 Directions for Determining the Starting and Concluding Tasks* (no date)

APPENDIX 3
Participation Guidelines for Alternate Assessment

The decision about a student's participation in assessment is made by the student's IEP team and documented in the IEP. To document that alternate assessment is appropriate for an individual student, the IEP team should review all important information about the student over multiple school years and multiple instructional settings (e.g., school, home, community) and determine that the student meets **all** of the following criteria:

- The student demonstrates a significant cognitive disability and adaptive skills, which result in performance that is substantially below grade-level achievement expectations even with the use of accommodations and modifications;

- The student accesses the state approved curriculum standards at less complex levels and with extensively modified instruction;

- The student has current adaptive skills requiring extensive direct instruction and practice in multiple settings to accomplish the application and transfer of skills necessary for application in school, work, home, and community environments;

- The student is unable to apply or use academic skills across natural settings when instructed solely or primarily through classroom instruction; and

- The student's inability to achieve the state grade level achievement expectations is not the result of excessive or extended absences or social, cultural, or economic differences.

Applicable Grades/Ages for Alternate Assessment Programs

South Carolina Readiness Assessment-Alternate Scoring (SCRA-Alt.)
The SCRA-Alternate Scoring should be used with students who meet all of the participation criteria for alternate assessment and whose age is commensurate with students in kindergarten and first grade (students who are 5 and 6 on September 1, 2006).

South Carolina Alternate Assessment (SC-Alt)
The SC-Alt should be administered to students who have been determined by the IEP team to meet all of the participation criteria for alternate assessment and who are age 8-13 or 15 on September 1, 2006.

APPENDIX 4
Alternate Assessment Alignment Pilot Study
Report to the South Carolina State Department of Education

Prepared by: Claudia Flowers, Diane Browder and Shawnee Wakeman,
University of North Carolina at Charlotte, and
Meagan Karvonen, Western Carolina University
April 2006

EXECUTIVE SUMMARY

This report details findings from an investigation of the alignment of South Carolina's alternate assessments in English language arts (ELA) and mathematics with other components of the educational system. The criteria used in this alignment study are being evaluated as part of the UNC Charlotte partnership in the *National Alternate Assessment Center* (NAAC). This report is organized by the seven criteria developed by a collaboration of content experts, special educators, and measurement experts at UNC Charlotte (Browder, Wakeman, Flowers, Rickleman, Pugalee, & Karvonen, 2006). While some of the alignment criteria are similar to other alignment methods (e.g., Webb, Surveys of Enacted Curriculum, and Achieve), some of the criteria (criteria 5-7) were designed specifically as value indicators for students with significant cognitive disabilities (see Table 1). An additional difference between this alignment protocol and other alignment methods is the examination of the targeted standards (i.e., standards intentionally selected for students with significant cognitive disabilities) and grade-level content standards. This summary describes how well the interpretation of state standards (Grade level and Measurement Guidelines), the alternate assessments (ELA-AA; Math-AA), and instruction (professional development manual and teacher survey about instruction) met the seven criteria for alignment.

Alignment Results by Criterion

Criterion 1: *The content is academic and includes the major domains/ strands of the content area as reflected in state and national standards (e.g., reading, math, science).*

> Outcome: The measurement guidelines were academic and reflected the major strands of reading and mathematics content (science was not reviewed) except for the omission of a focus on research skills in ELA. The alternate assessment also reflected the major strands of this content with a corresponding omission of research content. A few alternate assessment items were rated as nonacademic by the content experts because they did not fit any of the strands of ELA or mathematics content. These items were deleted from further alignment analysis. The professional development manual and teacher survey revealed a focus on the major strands of ELA and mathematics in instruction. Overall, this state system is aligned to academic content and meets criterion 1. We recommend either including content on Research in the measurement guidelines, alternate assessment, and professional development materials, or providing a rationale for why this ELA content strand is not considered relevant for this population.

Criterion 2: *The content is referenced to the student's assigned grade level (based on chronological age).*

> Outcome: For this second criterion, the focus was on alignment with the specific South Carolina curriculum standards for the content by grade bands in ELA and math. The measurement guidelines and alternate assessment items were aligned with the content

standards for the grade band. All categories of standards were represented except for the state standard on Research. The professional development manual reviewed was developed for an earlier era of alternate assessment and only contains information on linking to PK-2 standards. Overall, this state system is focused on grade level content standards in the measurement guidelines and alternate assessment. We recommend organizing professional development materials by grade bands.

Criterion 3: *The achievement expectation is linked to the grade level content, but differs in depth or complexity; it is not grade level achievement. It may focus on prerequisite skills or those learned at earlier grades, but with applications to the grade level content. When applied to state level alternate assessments, these priorities are accessible to IEP planning teams*.

Outcome: As would be expected for an alternate assessment based on alternate achievement standards, the measurement guidelines reflect levels of cognitive demand that are less complex than grade level achievement. The alternate assessment matches the depth of knowledge targeted by these measurement guidelines. For ELA measurement guidelines and alternate assessment, most items focused on reading. Math had a heavy emphasis on numbers and operations. At least 50% of the content standards under each academic domain had at least one MG or AA item except for the Research strand resulting in a 75% range-of-knowledge. Based on the teacher survey of instruction, in ELA, the majority of instructional emphasis was on reading, followed by communication and in math it was numbers and operations. In general, teachers identified a greater emphasis on the lower levels of cognitive demand as the highest performance expectation for the target student in 2005-06. Overall, this state has developed a system that targets achievement that is an alternative to grade level achievement. However, the balance across strands of content is weighted to one specific strand for both ELA and mathematics while reflecting some content in other strands. Currently, teachers report instruction that reflects similar emphasis by content area but with even lower levels of cognitive demand. Since the measurement guidelines and alternate assessments match in emphasis, these do align. We recommend some discussion about whether future work should focus on a wider range of knowledge for this population or maintain the current balance. We also recommend that professional development materials include ideas for teachers to increase the cognitive complexity reflected in instructional goals.

Criterion 4: *There is some differentiation in achievement across grade levels or grade bands*.

Outcome: This state uses the same alternate assessment across grade levels to show growth across grades. Our analysis revealed that there is a significant difference in the complexity of easier versus more difficult items in this assessment. The professional development materials do not yet indicate how to target increasing competence for a standard across grade levels/ grade bands.

Criterion 5: *The focus of achievement promotes access to the activities, materials, and settings typical of the grade level but with the accommodations, adaptations, and supports needed for individualization.*

Outcome: Because the state developed a single alternate assessment for use across grade levels, the goal was to utilize tasks that were applicable to all grades/ ages. Our analysis revealed that this goal was achieved as nearly all items were appropriate for either elementary or older students. In contrast, teachers reported that they adapted instructional materials primarily from grades K-2, even with students assigned to higher

grades. We recommend that the professional development materials contain information on how to adapt a grade level activity to students' current skill levels. The materials do include information on teaching in inclusive settings.

Criterion 6: *The focus of achievement maintains fidelity with the content of the original grade level standards (content centrality) and when possible, the specified performance (category of knowledge).*

Outcome: Overall content and performance centrality of the alternate assessment items to the measurement guidelines suggested a good quality of alignment. We recommend that professional development include guidelines for teachers on how to determine if an objective aligns to a state standard (e.g., see www.naacpartners.org resources for teachers on this topic.)

Criterion 7: *Multiple levels of access to the general curriculum are planned so that students with different levels of symbolic communication can demonstrate learning.*

Outcome: The alternate assessments contain items at all symbolic levels reflecting its accessibility for a wide range of students within this population. However, it is weighted heavily with items at the symbolic level. The professional development materials also contain examples at all symbolic levels although this specific terminology is not used. We recommend some state discussion of whether students below the symbolic level will/should be able to achieve proficiency on this alternate assessment with the number of items provided.

Overall Analysis of Alignment

This state has evidence supporting alignment for its measurement guidelines and alternate assessment based on all seven criteria. We conclude that overall this is an alternate assessment system that links to the grade level content. Some areas for consideration in further development of the system are noted above related to balance of content. We understood from the onset that the professional development materials reviewed have not been revised to reflect the current focus of the alternate assessment system. This was verified in our analysis as the materials currently support criteria 1, 2, and 7 but need additional material to address the remaining criteria. The information on instruction obtained from teachers was limited in both respondents and number of criteria addressed by the survey. However, it did suggest that the content of instruction roughly matched the alternate assessments while the cognitive complexity and grade level of adapted materials were not as well aligned.

APPENDIX 5

Technical Evaluation of Test Data From 2007 Administration: SC-Alt English Language Arts and Mathematics

South Carolina Alternate Assessment

Technical Evaluation of Test Data From Spring 2007 Administration:

SC-Alt English Language Arts and Math

A Report to the Educational Oversight Committee

Christine DiStefano
University of South Carolina
September 2007

South Carolina Alternate Assessment
Technical Evaluation of Test Data From Spring 2007 Administration:
English Language Arts and Math

Table of Contents

Description of the South Carolina Alternate Assessment Program

As part of South Carolina's state Accountability Program, students attending public schools take yearly standardized assessments to gauge student progress and relay information about school performance. Every student in the public schools is required to participate in the state testing program. This mandate also extends to students with cognitive disabilities. As stated on the SC Department of Education website (www.ed.sc.gov):

> "All students with disabilities must be included in statewide or district-wide assessments and if necessary, must have accommodations or modifications, or must participate in an alternate assessment."

An alternate assessment program has been developed to meet the needs of students with significant cognitive disabilities who are unable to participate in the general Palmetto Achievement Challenge Tests (PACT) or High School Assessment Program (HSAP) testing programs, even with accommodations and/or modifications. The SC assessment program for these students is the South Carolina Alternate Assessment (SC-Alt). The SC-Alt is an assessment for students with significant cognitive disabilities; these students are assessed against alternate achievement standards.

This report summarizes technical information from test data of the South Carolina Alternate Assessment (SC-Alt) in the areas of English Language Arts (ELA) and mathematics. Data for this report were collected as part of the 2007 operational administration of the SC-Alt. The Education Oversight Committee (EOC) supported the current study as part of its responsibilities listed in the Education Accountability Act of 1988:

> Section 59-18-320. (A) After the first statewide field test of the assessment program in each of the four academic areas, and after the field tests of the end of course assessments of benchmark courses, the Education Oversight Committee established in Section 59-6-10, will review the state assessment program and the course assessments for alignment with the state standards, level of difficulty and validity, and for the ability to differentiate levels of achievement, and will make recommendations for the needed changes, if any. The review will be provided to the State Board of Education, the State Department of Education, the Governor, the Senate Education Committee, and the House Education and Public Works Committee as soon as feasible after the field tests. The Department of Education will then report to the Education Oversight Committee no later than one month after receiving the reports on the changes made to the assessments to comply with the recommendations.

SC-Alt Population

The SC-Alt serves students with significant cognitive disabilities. Thus, students must meet eligibility criteria to be allowed to participate in the SC-Alt instead of the regular PACT or HSAP testing programs. To determine if a student is eligible for the SC-Alt, multiple sources of data are evaluated where the data are collected over a period of several years. Input from multiple sources and multiple time periods ensures that students who require additional assistance are eligible to take the SC-Alt.

The participation guidelines stated below are taken directly from the State Department of Education (SDE) website (www.ed.sc.gov):

The decision about a student's participation in assessment is made by the student's Individual Education Plan (IEP) team and documented in the IEP. To document that alternate assessment is appropriate for an individual student, the IEP team should review all important information about the student over multiple school years and multiple instructional settings (e.g., school, home, community) and determine that the student meets **all** of the following criteria:

- The student demonstrates a significant cognitive disability and adaptive skills, which result in performance that is substantially below grade-level achievement expectations even with the use of accommodations and modifications;
- The student accesses the state approved curriculum standards at less complex levels and with extensively modified instruction;
- The student has current adaptive skills requiring extensive direct instruction and practice in multiple settings to accomplish the application and transfer of skills necessary for application in school, work, home, and community environments;
- The student is unable to apply or use academic skills across natural settings when instructed solely or primarily through classroom instruction; and
- The student's inability to achieve the state grade level achievement expectations is not the result of excessive or extended absences or social, cultural, or economic differences.

Instead of following grade level requirements for testing, the SC-Alt is administered to students who have been determined by the IEP team to meet all of the participation criteria for alternate assessment and who are between the ages of 8-13 or are 15 years old as of September 1 of the current assessment year. The SC-Alt is organized into three test booklets based on grade level bands. The three forms are defined as:

- Elementary school form: covering grades 3 through 5 and appropriate for students between the ages of 8 - 10 as of September 1 of the current assessment year
- Middle school form: covering grades 6 through 8 and appropriate for students between the ages of 11 - 13 as of September 1 of the current assessment year
- High school form: covering grade 10 and appropriate for students 15 years of age as of September 1 of the current assessment year

The age bands were constructed for SC-Alt testing in lieu of following the students' stated grade level because students with significant cognitive disabilities may not make academic progress in the same manner as mainstream students.

SC-Alt: Test Development
Alignment of Test Content to Curriculum Standards

SC-Alt has been designed to meet all federal and state regulations concerning the test content. The content domains of the SC-Alt tests are aligned with alternative curriculum standards approved by the South Carolina State Board of Education. Alternative achievement standards are aligned with South Carolina achievement standards for mainstream students; however, the alternative achievement standards differ in the expectations of student performance as that they differ in complexity level. Curriculum standards for content areas covered by the SC-Alt are available on the SDE website (http://ed.sc.gov/agency/offices/assessment/programs/swd/SC-AltAssessmentStandardsandMeasurementGuidelines.html). The SC-Alt Assessment Standards and Measurement Guidelines were developed in compliance with the Individuals with Disabilities Education Act (IDEA) and the No Child Left Behind Act (NCLB) requirements that the alternate assessment must link to the grade-level content standards, although at less complex and prerequisite skill levels. More information about the link between the alternate curriculum standards and the SC-Alt test content is provided in the alignment study.

Test Design

SC-Alt replaces the previous alternate assessments, the PACT-Alt and the HSAP-Alt. The structure of the SC-Alt consists of a series of performance tasks in which students are required to demonstrate their understanding of the content. The SC-Alt tasks were developed by the testing contractor, American Institutes for Research (AIR), utilizing collaborative teams of experienced assessment writers with expertise in both the content areas and the learning characteristics of students with significant cognitive disabilities. The SC-Alt Assessment Standards and Measurement Guidelines provided the assessment teams with the ability to translate the standards into assessment tasks. The Content, Bias, and Accessibility Review Committee reviewed tasks prior to inclusion in the SC-Alt. The tasks were revised using input from small scale tryouts, focus groups discussions, and piloting and field testing to create the operational forms of the SC-Alt.

Each SC-Alt test form consists of twelve tasks. A task is a set of four to eight related activities or items and responses to the items provide evidence of what students know and can do in a given content area. Each test should have a sufficient number of items to provide a clear picture of student ability (Crocker & Algina, 1986) without overwhelming or fatiguing students.

While 12 tasks are included on each SC-Alt test form, the total number of items included on a test varies across the three grade band forms. For the operational forms of the 2007 spring administration of the SC-Alt, the numbers of items per form are provided below. Each form has a sufficient number of items included on each form to provide evidence of students' ability in a given content area.

Table 1. Number of Items on the South Carolina Alternate Assessment, ELA and Mathematics

Form	ELA	Mathematics
Elementary (Grades 3-5)	68	53
Middle School (Grades 6 – 8)	65	55
High School (Grade 10)	64	60
Total	197	168

Description of Testing Procedures

Given that a student meets the eligibility criteria for the SC-Alt and the correct grade level band is identified, teachers serve as test administrators for the SC-Alt. The test administrator administers the Student Placement Questionnaire (SPQ) to identify an appropriate starting position. The SPQ evaluates a student's ability and is used to determine an appropriate starting point within the test. This is done to avoid students being administered items that are too hard or too easy. Also, the process allows for an accurate assessment of the students' ability without overly fatiguing the student by exposure to unnecessary numbers of test items. Student fatigue is a concern given the dynamics of the SC-Alt population of students. Within a form, students are judged to have high, medium, or lower ability within the test band and the appropriate starting task is determined. Thus, students within the same grade level band may have different starting points within the same form, depending on the student's ability level. Given that students may have different starting points within the same instrument, students may, therefore, complete a different number of tasks. Additional detail about the SPQ and student placement is provided in the Test Administrators' Manual, which is available on the SC Department of Education website

(http://ed.sc.gov/agency/offices/assessment/programs/SWD/SouthCarolinaAlternateAssessmentSC-Alt.html).

SC-Alt test administrators undergo training to be familiar with the SPQ and how to interview students. Standardized training ensures that the teachers can gauge accurately an appropriate starting point. Additionally, the standardized training for all test administrators helps to ensure that the starting point judgments are fair and unbiased.

Each item on the SC-Alt has a point worth which may vary from one point to four points, depending on the complexity of the task to be performed. The test administrator scores the SC-Alt assessment as it is administered. To ensure scoring fidelity and scoring standardization across the state, training is required for all teachers who will administer the SC-Alt assessment. Standardized training for every test administrator helps to ensure appropriately administered and scored assessments. Proper test administration and scoring supports the validity of the SC-Alt results used for Adequate Yearly Progress (AYP) ratings and school report card ratings.

Sample Size

The SC-Alt is a specialized instrument, where students must meet pre-specified conditions to be eligible to take this test. The estimated number of students taking the SC-Alt is approximately 0.05% of the student population in SC schools (SC-Alt Technical Manual, March 16, 2007). The SC-Alt Technical Manual states that students with three primary disability designations

accounted for approximately 80% of the participants: trainable mentally disabled students (51.2%), autistic students (14.6%), and profoundly mentally disabled students (14.0%).

The number of students tested in the spring 2007 administration of the SC-Alt assessment was reported in the July, 2007 Summary Tables provided to the SC-Alt Technical Committee (AIR Technical Team, July 2007). Student sample sizes for the spring 2007 administration of the SC-Alt are provided in Table 2. Test data from these operational samples was used to compute the item statistics evaluated in the current report. The number of students involved with the spring 2007 SC-Alt administration is acceptable for students in the Elementary and Middle School grade level bands. It is recognized that the sample size for the High School grade band is lower than desired; however, this sample size represents disabled students within the grade band who were eligible to take the SC-Alt.

Table 2. Number of Students Tested, 2007 South Carolina Alternate Assessment ELA and Mathematics Assessment

Form	ELA	Mathematics
Elementary (Grades 3-5)	1076	1065
Middle School (Grades 6 – 8)	989	982
High School (Grade 10)	335	339

Data Analysis Procedures

SC-Alt item statistics were calculated by the SDE/AIR and delivered to the EOC for evaluation. EOC staff provided the SDE data sets to this author. Data sets contained statistical information for the SC-Alt ELA and Mathematics Fall 2007 operational administrations. Item statistics were calculated using Classical Test Theory (CTT) techniques and Item Response Theory (IRT) techniques where the Rasch model (i.e., one parameter item response theory model) was used. For the technical report, summaries of item statistics (difficulty, average point biserial values) and psychometric characteristics (e.g., Differential Item Functioning, Rasch ability estimates) were summarized for SC-Alt ELA and mathematics operational forms. It is noted that this technical report consists of evaluation and interpretation of the dataset indices provided to the EOC. Besides calculation of summary statistics (e.g., mean values, standard deviations), no additional estimation procedures (e.g., equating, ability estimates) were conducted. This report is arranged into three sections: a) summary of classical test theory indices, b) summary of item response theory indices, and c) investigation of impact.

Section A: Summary of Classical Test Theory Indices

Two Classical Test Theory (CTT) indices were included on the dataset: item difficulty and adjusted point-biserial. The item difficulty (p) may be defined as the proportion of students out of the total number of examinees answering an item correctly. Higher p values indicate easier items (i.e., a greater number of students selected the correct answer) and low p-values indicate more difficult items. Items which are too difficult or, conversely, too easy, do not differentiate between low performing and high performing students. A difficulty value of p = .5 provides the highest level of differentiation between students (Crocker & Algina, 1986).

The adjusted point biserial *r* is a measure of association indicating how well an item discriminates between high performing and low performing students. The value is calculated as the correlation between item scores (correct/incorrect) and the total score, with the item in question removed from the total score. The normal range of point biserial scores for items is –1 to +1, with higher values indicating that the item discriminates well between high and low performing students (Crocker & Algina, 1986). Values of the point biserial may be positive, meaning that the item is discriminating appropriately, or negative, indicating that the item is not discriminating as intended. Values that are close to zero or negative may indicate a flawed item. A value of zero means that there is no discrimination between high and low ability test takers; negative values indicate the tendency for high ability students to answer incorrectly and low ability students to answer correctly. A high point-biserial coefficient means that students selecting the correct response are students with higher total scores, and students selecting incorrect responses to an item have lower total scores, meaning the item can discriminate between low-performing examinees and high-performing examinees.

CTT Difficulty

Table 3 provides summary statistics for the difficulty values by SC-Alt Test form and age band and content area. Mean values across the ELA forms were at least p= .63, meaning that, on average, students answered 63% of the SC-Alt ELA items correctly. Minimum and maximum p-values showed a range of item difficulty values, ranging from a minimum value of p = .247 (illustrating a difficult item) to p=. 875 (illustrating a relatively easy item).

Item difficulty values were reviewed to determine the number of ELA items per form that were challenging for students, where p < .50. On the Elementary Form, 11 of the ELA 68 items (16%) had a p-value less than or equal to .50, 13 of 65 items (20%) on the Middle School form were challenging for students, and 15 of 64 items (23%) on the High School form were challenging. Thus, the majority of the SC-Alt ELA items were relatively easy for the population of students.

Mean values across the SC-Alt mathematics forms were at least p= .62, meaning that, on average, students answered at least 62% of the math items correctly. Minimum and maximum p-values in mathematics reported item difficulty values, ranging from p = .333 (relatively difficult for students) to p=. 875 (relatively easy items). Again, item difficulty values were reviewed to determine the number of items per form that were challenging for students, where p < .50. On the Elementary Form, 9 of the 53 items (17%) had a p-value less than or equal to .50, there were 13 of 55 items (24%) on the Middle School form with p-values less than .50, and 13 of 60 items (22%) on the High School form.

For the SC-Alt tests in ELA and mathematics, the information showed that the tests became slightly harder for students as the age band increased from Elementary to High School. This is not unusual, given that the content also increases in difficulty. Overall, for SC-Alt ELA and Mathematics content, difficulty values are slightly easier than expected (with p=.50 set as the midpoint for difficulty) Difficulty values are within an acceptable range, especially given the nature of the population, the use of the SPQ to pinpoint the appropriate student starting point, and the purpose of the SC-Alt instrument.

Table 3. CTT Difficulty Values, by Form

Form and Age Band	Number of Items	Mean Difficulty	Standard Deviation	Minimum Difficulty	Maximum Difficulty
ELA					
Elementary	68	.658	.142	.367	.859
Middle	65	.660	.154	.356	.875
High School	64	.633	.162	.247	.844
Mathematics					
Elementary	53	.631	.140	.367	.875
Middle	55	.629	.146	.333	.836
High School	60	.615	.123	.346	.817

CTT Discrimination

Table 4 provides summary statistics for the adjusted point biserial values for the SC-Alt ELA and Mathematics tests. Mean values across the SC-Alt ELA forms was at least r_{pb} = 0.41, illustrating that the set of tests are moderately discriminating. The average value means that, generally, SC-Alt students with lower total test scores chose incorrect responses and higher ability students chose correct responses. However, the r_{pb} is lower than .5, showing some inconsistencies. As seen by the mean point biserial value by form, the SC-Alt forms were roughly equivalent in their ability to discriminate between higher and lower ability students; no one form discriminated significantly better (worse) than the other SC-Alt ELA forms.

Item point biserial values were reviewed to determine the number of items per form that were able to discriminate between students of high and low ability students, where r_{pb} was greater than or equal to .50. ELA SC-Alt items were discriminating between students of different ability levels. On the Elementary Form, 18 of the 68 items (27%) had a adjusted point biserial values greater than or equal to .50, 27 of 65 (42%) of Middle School items reported r_{pb} greater than or equal to .50, and 28 of 64 items (44%) on the High School form were above .50. These values show that the tests are increasingly discriminating as the grade level band increases

A similar pattern was seen for the SC-Alt Math forms, where the mean point biserial value was at least .40, indicating a moderate level of discrimination. Item point biserial values were reviewed to determine the number of items per form that were able to discriminate between students of high and low ability students above the midpoint value (r_{pb} > .50). On the Elementary Form, 17 of the 53 items (32%) had adjusted point biserial values less than or equal to .50 and 20 of 60 items (33%) on the High School form – meaning that roughly a third of the items on

these forms were very good for discriminating between higher and lower ability students. The SC-Alt mathematics middle school form was more difficult for students with 29 of 55 items (53%) on the form yielding point biserial-values greater than .50.

Over both the SC-Alt ELA and Mathematics forms, the items are able to differentiate between students of higher and lower ability. The items are performing adequately to judge student knowledge. The discrimination information is thought to be appropriate given the requirements of the SC-Alt.

Table 4. Adjusted Point Biserial Values, by Form

Form and Age Band	Number of Items	Mean r_{pb}	Standard Deviation	Minimum r_{pb}	Maximum r_{pb}
ELA					
Elementary	68	.414	.157	.162	.759
Middle	65	.412	.140	.114	.695
High School	64	.479	.111	.201	.721
Mathematics					
Elementary	53	.406	.142	.091	.675
Middle	55	.493	.097	.288	.668
High School	60	.448	.100	.214	.675

Section B: Summary of Item Response Theory Indices

IRT models are represented by statistical functions which relate person and item characteristics to the probability of choosing a correct item response. IRT uses a model based approach to: estimate item parameters, determine how well the data fit the model, and to investigate the psychometric properties of items and tests (Baker, 2001). A one-parameter IRT model, the Rasch model, was applied to the SC-Alt operational test data to obtain item parameters and fit information. Three IRT indices were included on the dataset: Infit and Outfit fit statistics, and Rasch item difficulty. Items were flagged if they exhibited differential performance for one subgroup compared to another. Items exhibiting differential item functioning (DIF) may be easier or more difficult for one demographic group compared to another, and should be examined to rule out the possibility that they may bias the test results.

A characteristic of the Rasch model is that all items are thought to have the same item discrimination, but varying levels of item difficulty. The difficulty parameter is defined as the point on the ability scale at which the probability of correct response to the item is .5, where the slope of the Rasch curve is at a maximum. Typical values are within the range –3 < = *difficulty* < = +3. (Baker, 2001). Item difficulty parameters can be interpreted relative to ability level. As stated in Baker (2001, p. 34-35) " an item whose difficulty is –1 functions among lower ability examinees while an item with a difficulty value of +1 does best to distinguish between examinees functioning at higher ability levels."
Both Infit and Outfit are fit statistics, which indicate in the Rasch context how accurately the data fit to the Rasch model. As stated by Bond & Fox (2001):

> Outfit statistics have more emphasis on unexpected responses far from a person's or item's measure. Infit statistics place more emphasis on unexpected responses near a person's or item's measure.

Stated another way by the Winsteps user's manual (Linacre, 2006, http://www.winsteps.com/winman/diagnosingmisfit.htm**)**

> Outfit measures are more sensitive to unexpected observations by persons on items that are relatively very easy or very hard for them (and vice-versa). Infit measures are more sensitive to unexpected patterns of observations by persons on items that are roughly targeted on them (and vice-versa).

Infit and outfit values can be reported as unstandardized values, standardized values, or mean square values. To be consistent with the infit/outfit item flag information, mean square values will be discussed. Mean square values are computed as the Rasch model chi-square statistic divided by its degrees of freedom (http://www.winsteps.com/winman/diagnosingmisfit.htm). Expected values for the mean squares should approximate 1.0. Values greater than 1 (underfit) indicate unmodeled noise or other source of variance in the data and may degrade measurement. Values less than 1 (overfit) indicate that the model predicts the data too well, and may cause summary statistics to report inflated values.

IRT Difficulty Indices

Rasch item parameters provide a modern test theory perspective of item difficulty. Most difficulty values for the SC-Alt operational items are functioning slightly below the mean ability level of 0 for both ELA and Math. The information shows that the items are functioning best for students with slightly lower than average ability levels in this population of students. The SC-Alt High School forms are slightly harder for students, as shown by mean difficulty values closer to 0. For mathematics, the Elementary test is the least difficult, with mean difficulty values at -.58. The SC-Alt Middle School and High School forms are more difficult, with difficulty values of -.46 and -.32, respectively. Overall, the tests are increasing in difficulty as the grade band increases.

Difficulty values are negative for the SC-Alt ELA and mathematics forms, meaning that the items function best with students who have lower than average ability. Calculations showed that mean Rasch difficulty values for each form were smaller than the median Rasch difficulty values, reflecting negative skewness in the distribution of IRT difficulty scores. For ELA and mathematics item statistics, difficulty values appear to be within acceptable ranges. Standard deviation values are above .55, suggesting that the assessments included a reasonable range of item difficulties. Table 5 provides summary statistics across the SC-Alt ELA and mathematics forms.

Table 5. IRT Based Difficulty Values, by Form

Form and Age Band	Number of Items	Mean Difficulty	Standard Deviation	Minimum Difficulty	Maximum Difficulty
ELA					
Elementary	68	-.57	.75	-2.26	1.04
Middle	65	-.56	.72	-2.26	1.04
High School	64	-.19	.55	-1.73	.98
Mathematics					
Elementary	53	-.58	.66	-2.26	.87
Middle	55	-.46	.70	-2.26	.87
High School	60	-.32	.68	-2.26	1.25

Infit and Outfit Measures

Tables 6 and 7 below provide the mean square values for infit and outfit. For both infit and outfit mean square values, mean values suggest adequate fit. All items appear to have average levels of infit/outfit close to the expected value of 1. This indicates that the Rasch model provides an acceptable fit to the operational test data for the SC-Alt ELA and SC-Alt mathematics forms.

Table 6. Average Standardized Infit Values, by Form

Operational Form and Age Band	Number of Items	Mean Infit	Standard Deviation	Minimum Infit	Maximum Infit
ELA					
Elementary	68	1.00	.16	.72	1.49
Middle	65	1.00	.17	.78	1.49
High School	64	1.01	.16	.74	1.51
Mathematics					
Elementary	53	1.00	.12	.75	1.39
Middle	55	1.00	.12	.74	1.39
High School	60	1.04	.15	.83	1.74

Table 7. Average Standardized Outfit values, by Form

Operational Form and Age Band	Number of Items	Mean Outfit	Standard Deviation	Minimum Outfit	Maximum Outfit
ELA					
Elementary	68	1.00	.26	.54	1.83
Middle	65	.99	.29	.44	1.83
High School	64	.97	.25	.49	1.66
Mathematics					
Elementary	53	.98	.21	.57	1.60
Middle	55	1.00	.18	.58	1.60
High School	60	1.05	.23	.67	2.09

Differential Item Functioning

Items on the SC-Alt ELA and mathematics subtests were examined for differential item functioning (DIF). DIF analyses identify items that do not perform equally across subgroups of the SC-Alt population. Comparisons were made between sex groups (male and female students) and racial groups (Black and Caucasian students). If many items exhibit DIF, the test may give one group an unfair advantage (disadvantage) over other test takers. Here, DIF is discussed in general terms. Specific items that are exhibiting DIF are named in the Item Flags section.

For the SC-Alt ELA tests, two items reported differential item functioning at severe levels on the middle school form and six items showed problems on the high school form. No items exhibiting DIF were found on the SC-Alt ELA Elementary form. For the two items reporting DIF on the middle school form, both items were cited for differential performance based upon students' sex. On the high school form, all six items yielded differential functioning depending on student race.

These items could be reviewed for problems (such as content, wording, etc.) to try to eliminate DIF in future administrations of the SC-Alt ELA tests.

For the SC-Alt mathematics tests, only the high school form reported items that exhibited DIF. There were no items that exhibited DIF on the SC-Alt Middle School or Elementary forms. For the items showing DIF on the SC-Alt high school mathematics form, four of the items reported differential performance between sexes, seven items reported differential performance between racial groups, and two items exhibited DIF for both sex and race subgroups. As with the SC-Alt ELA items that showed evidence of DIF, items exhibiting DIF on the SC-Alt mathematics forms may be reviewed to try to eliminate DIF in future administrations of the SC-Alt ELA tests. This suggestion is more pressing for items that exhibit DIF across both sex and racial groups. It is also recognized that the SC-Alt high school sample size is the smallest of the three forms. The small sample size, and even smaller subgroup sample sizes, may exert undue influence on the item statistics.

Item Flags

A flagged item suggests that the performance may be problematic and the item may need a closer inspection. Items were flagged by the SDE for a variety of performance indicators. While many flags could be noted, the six flags that were present in the SC-Alt dataset are described below. Descriptions of the item flags were taken from the SDE/AIR data codebook:

- Difficulty flags indicated items that were excessively hard ($p < .30$) or too easy ($p > .95$)
- Point biserial flags for low biserial correlations ($r_{pb} < .20$) meaning that the item was not discriminating between students of higher and lower ability levels.
- Differential item functioning (DIF) illustrates that an item may be easier or more difficult for one demographic group compared to another
- Fit if infit <.7 or infit >1.3 or outfit <.7 or outfit >1.3
- Omit flags suggest that the item's omit rate is too large, i.e., >.05, meaning that roughly 5% of the students' omitted this item
- CRT flagged items were those flagged if the mean total test score of students in a score point category was lower than the mean total test score of students in the next lowest score point category. For example, if students who received 3 points on a constructed response item scored, on average, lower on the total test than students who received 2 points on the item, the item would be flagged. This situation may indicate that the scoring rubric is flawed.

For the SC-Alt database, all item characteristics were examined. Items were flagged for violating one rule or a combination of the rules.

Information concerning flagged items on the SC-Alt ELA tests is provided in Table 8. As Table 8 shows that 53 out of 197 ELA items were flagged for various problems. Stated another way, approximately 26.9% of the set of ELA items yielded item statistics which were outside of the stated bounds. The percent of items showing problems was 13 of 68 (19.1%) of items flagged on the elementary form, 18 of 65 (27.7%) of items flagged on the middle school form, 22 of 64

(34.4%) of items flagged on the high school form. The number of flags observed is somewhat surprising given that the test has already undergone item screening, item revision, and field-testing procedures. However, the majority of flags were given infit/outfit statistics being outside of stated boundaries. The information suggests that the model is not predicting the data accurately, where unmodeled variance may be present. This variance could be due to other sources such as individual student characteristics, disability type, or even student fatigue.

Differential item functioning (DIF) is a more serious flag. As discussed earlier, items exhibiting DIF were found on the middle and high school SC-Alt ELA forms. While DIF indicates differential performance, there are relatively few items out of the entire test that exhibit DIF. Also, it is noted that there are relatively few students in the entire SC-Alt population as compared to the mainstream population of students. Depending on the size of the subgroup, if high numbers of students from a subgroup have problems with an item small sample size could lead to misrepresentation of an item's performance.

Table 8. Item Flags, SC-Alt English Language Arts Tests

Flags				
Form	No. Of occurrences	Percent Flagged	Type of Flag(s)	Item numbers
ELA	**53**			
Elementary	13			
	1		r_{pb}	60
	7		Fit	24, 41, 49, 58, 63, 64, 66
	1		Omit & Fit	7
	4		r_{pb} & Fit	57, 61, 65, 67
Middle	18			
	2		r_{pb} & Fit	54, 64
	13		Fit	52 49 32 1 40 36 48 62 34 8 15 35 33
	2		DIF	3 13
	1		Omit	2
High School	22			
	2		Crt	13 52
	5		DIF	56 18 32 11 3
	13		FIT	58 54 25 22 33 62 46 55 10 1 48 29 53
	1		CRT & Fit	63
	1		Omit & DIF	6

Information concerning flagged items on the SC-Alt mathematics tests is provided in Table 9. Across the three forms, 35 out of 168 mathematics items were flagged (20.8%).The percent of items showing problems was low by form with 14 of 53 (26.4%) of items flagged on the elementary form, 5 of 55 (9.1%) of items flagged on the middle school form, 16 of 60 (26.7%) of items flagged on the high school form. Again, the numbers of SC-Alt mathematics items flagged was somewhat unexpected given that the mathematics has already undergone item screening, item revision, and field-testing procedures.

Overall, most SC-Alt mathematics items were flagged for evidence of infit and/or outfit statistics. This means that the items are not performing adequately and are producing scores that may be unexpected. Again, while this flag is present, it is not overly serious. Other flags, such as point

biserial, CRT, and Omitted items, were observed, but these flags made up a relatively small percent of the total of flagged items.

The SC-Alt mathematics High School form showed the most flags, roughly 27% of the test items cited. The most disconcerting information here is the number of items showing evidence of DIF, meaning that the items were performing differently for different subgroups of test takers. Nine of the 16 flagged items showed presence of DIF. These items may be re-examined to determine if the amount of differential functioning is high enough to bias the test for different groups of SC-Alt students.

Table 9. Item Flags, SC-Alt Mathematics Tests

Flags				
Form	No. of occurrences	Percent Flagged	Type of Flag(s)	Item numbers
Mathematics	**35**	**20.8**		
Elementary	**14**	**26.4**		
	7		Fit	1, 27, 30, 38, 44, 45, 46
	4		r_{pb}	47, 48, 49, 52
	1		r_{pb} & Fit	53
	2		Omit	3, 5
Middle	5	9.1		
	4		Fit	1, 26, 31, 55
	1		CRT	47
High School	16	26.7		
	5		Fit	1, 28, 39, 53, 58
	7		DIF	3, 30, 41, 44, 47, 50, 59
	1		CRT	57
	1		Fit & CRT	24
	1		Fit & DIF	60
	1		CRT & DIF	9

Section C: Estimates of Impact

To judge impact of the SC-Alt, the assessments should be able to categorize students into different ability levels, according to the amount of knowledge students possess in a given content area. The SC-Alt assessments categorize students into one of four achievement levels. The levels are named 1, 2, 3, and 4, where level 1 represents the lowest achievement level and level 4 represents the highest achievement level on the SC-Alt. The descriptions of the SC-Alt achievement levels were created by the SDE and AIR and provide a detailed assessment of student competencies and skills that students must demonstrate to be "graded" at a specific level of performance. Performance descriptors vary by content area and grade level band. While detailed information about the achievement level descriptors is provided in the SC-Alt Standard Setting Technical Report (AIR, September, 2007), a generic description of the achievement levels is provided in Table 10. The generic description shows the increasing performance and knowledge requirements for the SC-Alt in ELA and mathematics as the achievement level increases from level 1 to 4.

Table 10. Generic Description of SC-Alt Assessment Achievement Levels

Level	Generic description of SC-Alt Assessment Achievement Levels
Level 1	Students performing at level 1 may demonstrate emerging academic skills and competencies in ELA (mathematics).
Level 2	Students performing at level 2 demonstrate foundational academic skills and competencies in ELA (mathematics).
Level 3	Students performing at level 3 demonstrate increasing academic skills and competencies in ELA (mathematics).
Level 4	Students performing at level 4 demonstrate and apply academic skills and competencies in ELA (mathematics).

AIR, under contract to the SC SDE, held a workshop to recommend performance standards for the SC-Alt assessments. The workshops were held June 25-27, 2007 and involved 105 educators and non-educators (e.g., parents, curriculum specialists) from across the state. The panel recommended standards to categorize students into levels 2, 3, and 4 on the SC-Alt assessments. The standards were translated into cut points on the SC-Alt tests by AIR.

Using the information from the cut scores, it is of interest to estimate the impact of the SC-Alt assessments by evaluating average student ability estimates for the SC-Alt ELA and mathematics tests. It is noted that the information evaluated in Table 11 was taken directly from AIR technical documentation. At the time of this report (September 5, 2007), impact results for the spring 2007 administration of the SC-Alt have not been published by the SDE. The information presented in Table 11 allow for an initial investigation of impact; however, additional impact data may be examined and evaluated at a future date.

Table 11 shows the range of ability estimates for each performance level on the SC-Alt ELA and mathematics tests. Ability estimates range from negative infinity to positive infinity, thus no minimum for level 1 and maximum for level 4 are needed in the table. As expected, the higher the performance level, the higher the students' estimated ability. Ability estimates were lower than average (i.e., ability = 0) only for the lowest performance levels, levels 1 and 2. Overall, the SC-Alt ability estimates appear to be within adequate ranges and the categorization of students into different performance levels allows for differentiation of students at different ability levels.

Table 11. Estimates of Impact by Grade Range, SC Alt Assessment

	Level	Cut Scale Score	Minimum Ability Estimate	Maximum Ability Estimate
ELA Grade 3-5	Level 1	--	*	-1.21
	Level 2	403	-1.20	-0.03
	Level 3	466	-0.02	0.66
	Level 4	491	0.67	*
ELA Grade 6-8	Level 1	--	*	-0.89
	Level 2	417	-0.88	0.18
	Level 3	473	0.19	0.79
	Level 4	501	0.80	*
ELA Grade 10	Level 1	--	*	-0.94
	Level 2	429	-0.93	-0.03
	Level 3	478	-0.02	0.66
	Level 4	503	0.67	*
Math Grade 3-5	Level 1	--	*	-1.07
	Level 2	423	-1.06	0.08
	Level 3	476	0.09	0.73
	Level 4	526	0.74	*
Math Grade 6-8	Level 1	--	*	-1.01
	Level 2	425	-1.00	0.08
	Level 3	476	0.09	0.95
	Level 4	529	0.96	*
Math Grade 10	Level 1	--	*	-0.93
	Level 2	434	-0.92	-0.28
	Level 3	476	-0.27	0.50
	Level 4	528	0.51	*

Notes: No cut score is needed to categorize students into Level 1.

Summary and Recommendations

This report summarizes the results from the spring 2007 operational administration of the South Carolina Alternative (SC-Alt) assessments. The SC-Alt is geared towards students with cognitive deficiencies who are unable to take the regular state assessments, even with modifications. The Education Oversight Committee (EOC) supported the current study as part of its responsibilities listed in the Education Accountability Act of 1988. This study reviewed item and form data from the English Language Arts (ELA) and Mathematics forms administered spring 2007. Test information was presented for three age bands: Elementary (3-5), Middle school Form (6-8) and High School (10). Indices of Classical Test Theory (CTT) and Item Response Theory (IRT) were interpreted by form and subject area. Based on the results, the following evaluations and recommendations are provided.

A strength of the SC-Alt assessment battery is the interrelationship between the components of the assessment system. The SC-Alt tests were revised to include performance tasks, which were thought to better estimate the knowledge and ability of students with significant cognitive disabilities. Also, multiple sources of evidence collected over a long period of time are evaluated to determine if a student is eligible for the SC-Alt instead of the state's mainstream testing program. Using a variety of evidence collected from multiple sources helps ensure that students in need of the alternative program are eligible for the assessment. This helps to provide an accurate reflection of the population of cognitively disabled students across the state. Finally, the standardized training given to test administrators for student placement on the test and scoring of responses helps to ensure that the scores obtained from the SC-Alt are valid measures of student ability and can be trusted to make inferences of student ability.

Overall, the SC-Alt ELA and mathematics tests appear to be functioning adequately for the three different grade bands studied. It was noted that the sample size used to calculate CTT and IRT statistics with the high school test (Grade 10) was lower than the sample sizes used in the other two tests. However, the SC-Alt population is a special needs population where relatively few students across the state fall into this category (estimate of .5% of SC public school students).

The ELA and mathematics forms generally reported CTT and IRT item statistics which were within acceptable ranges. The tests are of increasing difficulty and can be used to differentiate students based on ability. The Student Placement Questionnaire helps ensure that students gain an optimal starting place to measure their content knowledge. Both CTT and IRT estimates of difficulty reported that the test was performing adequately; for a given form, students answered approximately 60% of items correctly. Also, the test reported acceptable levels of discrimination, indicating that the ELA and mathematics tests were able to distinguish between high and low ability students. The test is not maximally discriminating; however, this may be acceptable given the requirements of the SC-Alt testing program.

In terms of item performance, many items were flagged due to problematic item statistics. It is noted that the majority of flags were given for infit/outfit IRT measures rather than something more serious. However, roughly one third of the items on a given form were flagged for some sort of problematic behavior. It is recommended that the items be reviewed with future operational administrations of the test. Over 15 items showed significant Differential Item Functioning (DIF) between subgroups of SC-Alt students on either the ELA or mathematics forms. These items should be investigated further to ensure that items do not function differently for subgroups of students. It is recommended that these items be reviewed in future

administrations of the SC-Alt examination. If many items are still problematic, the items may be reviewed to see if wording problems are apparent or if increasing item clarity may improve item performance. Finally, because impact data were not available at the time of this report, future evaluations of SC-Alt test data should evaluate estimates of impact to ensure that the estimates of student ability are in agreement with the objectives of the SC-Alt. This should include an evaluation of the percentage of students classified into each performance level (i.e., level 1 through level 4), review of ability estimates by performance level, and review of the grading rubrics used to categorize student performance.

In summary, the technical information suggested that the SC-Alt ELA and mathematics forms were performing acceptably. Selected items showing DIF and performance rubrics for ELA were suggested for review with data from future operational administrations of the tests. Overall, the SC-Alt appears to perform effectively to assess South Carolina's students with significant cognitive disabilities.

Reference List

Baker, Frank (2001). *The Basics of Item Response Theory.* ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD.

Bond, T. G., and Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Lawrence Erlbaum Associates: Mahwah, NJ.

Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Harcourt, Brace & Jovanovich, Inc: Orlando, FL.

Linacre, J. M. (2006) Winsteps Rasch measurement computer program. Chicago: Winsteps.com. Information retrieved from: http://www.winsteps.com, August 28, 2007.

South Carolina Department of Education (2007). Information retrieved from http://www.ed.sc.gov, August 28, 2007.

The Education Oversight Committee does not discriminate on the basis of race, color, national origin, religion, sex, or handicap in its practices relating to employment or establishment and administration of its programs and initiatives. Inquiries regarding employment, programs and initiatives of the Committee should be directed to the Executive Director 803.734.6148.